RAPID APPROXIMATION OF SERVOMECHANISM TRANSIENT RESPONSE

WALTER BUEL PATTON
and
WILLIAM BENJAMIN ABBOTT, III

RAPID APPROXIMATION OF SERVOMECHANISM TRANSIENT RESPONSE

by

MAJOR WALTER BUEL PATTON, USMC

B.S.Ch.E., Auburn University
(1943)
B.S.E.E., U.S. Naval Postgraduate School
(1960)

and

LIEUTENANT WILLIAM BENJAMIN ABBOTT III, USN

B.M.E., Georgia Institute of Technology
(1952)
B.S.E.E., U.S. Naval Postgraduate School
(1960)

SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE
DEGREE OF MASTER OF
SCIENCE
at the
MASSACHUSETTS INSTITUTE OF
TECHNOLOGY
June, 1961

RAPID APPROXIMATION OF SERVOMECHANISM TRANSIENT RESPONSE

Major Walter Buel Patton, USMC
Lieutenant William Benjamin Abbott III, USN

Submitted to the Department of Aeronautics and Astronautics on May 1, 1961 in partial fulfillment of the requirements for the degree of Master of Science.

ABSTRACT

The purpose of this thesis is to develop a means by which the transient response of a servomechanism may be rapidly evaluated. The approach consists of developing a method of rapidly determining the value of several transient response parameters so that the system transient response may be sketched. The parameters selected for this purpose are time and magnitude of peak overshoot, rise time and settling time.

To permit evaluation of these parameters systems are classified according to the number and type of singularities in their closed loop performance functions. Systems considered are underdamped systems with two or three singularities. For each class of system, loci of constant values of overshoot and constant values of the three time parameters expressed in terms of an artificial time are mapped onto a complex plane. Parameter loci for systems with only two singularities are mapped directly onto the s-plane. Loci for systems with three singularities are mapped onto an auxiliary complex plane.

The values of the four transient response parameters for any system of a particular class can be read from the families of loci for that class. The magnitude of overshoot can be used directly and the three time parameters can be easily converted into real time. The initial slope of the transient response and the slope at the time of peak overshoot are either known or easily determined. Also the transient oscillating frequency is obtainable directly from the system closed loop performance function. Using this information the transient response of a particular system can be sketched with considerable accuracy. This sketch can then be used as a basis for rapid system evaluation.

Thesis Supervisor: Robert K. Mueller

Title: Associate Professor of Aeronautics and Astronautics

ACKNOWLEDGMENTS

The authors wish to thank Professor Robert K. Mueller, who acted as their thesis advisor, for the encouragement and counsel he afforded them in friendly fashion during the course of their work.

The idea of mapping time domain quantities onto a complex plane first arose during discussions between one of the authors and Lieutenant Charles A. Phillips, USN, a classmate at the U.S. Naval Postgraduate School.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

Chapter I

INTRODUCTION

In the past two decades the importance of servomechanisms and automatic control systems in general has increased tremendously. Accordingly, much effort has been expended in developing methods of analysis and design. The result is an abundance of ideas and techniques available to the designer today. Most of these can be found in any one of a number of current automatic control system textbooks.

Basically there are two approaches to analysis and design, the frequency response approach and the time domain approach. Either of these techniques may be used with some facility in designing systems to meet both time domain and frequency response specifications. However, if time domain specifications are to be met, a time domain analysis must eventually be made. This is true regardless of which technique is used in the actual design.

The usual time domain specification is one involving transient response, that is, the time response of the system to a suddenly applied input signal. The most common input signal used is the unit step. Hereafter we will use the term <u>transient response</u> exclusively to mean the time response of a system to a unit step input signal with all initial conditions equal to zero.

The problem of obtaining the transient response of servomechanisms is one of considerable magnitude. The transient response of linear second order systems has been studied extensively and is well known. The transient responses of higher order systems and second order systems that have been complicated by addition of a zero are more difficult to obtain and are less well known. The exact transient response of any linear system can always be obtained by factoring the characteristic equation, solving for the output as a function of time and plotting this function. However, this procedure involves considerable labor and is quite time consuming.

One method of attempting to avoid such a calculation is to consider the complex roots of the characteristic equation to be dominant and use a second order approximation of the more complicated system. As the value of the real roots and zeros become smaller with respect to the undamped natural frequency of the system this approximation ceases to be valid. When the second order approximation ceases to be valid another approximation developed by Chu (1) may be employed. This method reduces the labor involved in obtaining the transient response by applying corrections to the second order equations for zeros and other roots of a system. Unfortunately this approximation also becomes invalid as the values of the real roots and zeros become even smaller.

If the frequency response approach is used, obtaining the transient response is usually more difficult because the system

characteristic equation is not known in factored form. In this case graphical integration schemes such as one developed by Floyd (2) are sometimes used to obtain an approximate transient response. These methods also require considerable labor and time.

In this thesis we restrict the discussion to linear systems which exhibit the property of unity feedback, that is, systems for which the steady state output becomes unity when a unit step input signal is applied. Notice that absolute stability is implied.

With this restriction we can make some general observations about the transient response of any such system.

a. At zero time the value of the output is zero.

b. At a sufficiently large time the value of the output becomes unity.

c. For critically damped and overdamped systems the output increases monotonically with time until it reaches its steady state value of unity. This case is of little practical interest.

d. A more common and more practical case is that of an underdamped system. In this case the output rises to a value greater than unity, and may or may not oscillate about unity before assuming its steady state value.

Based on the above assumptions and observations we may define the following transient response parameters:

a. Peak Overshoot - M_p The amount by which the maximum value of the output exceeds the steady state value.

b. Time of Peak Overshoot - t_p The time at which the maximum value of output occurs.

c. Rise Time - t_r The time at which the value of the output is unity for the first time.

d. Settling Time - t_s The mathematical expression for the transient response of systems which have only one pair of complex roots may consist of three parts. These are a constant equal to the steady state value of output, decaying exponential terms, and a damped sinusoidal term. The settling time is the time at which the sum of the error due to the exponential terms plus the height of the envelope of the damped sinusoid is equal to two per cent of the steady state output value.

A sketch of a typical transient response with the above defined parameters indicated is shown in Figure I-1.

If the four transient response parameters defined above and the transient oscillating frequency of a system are known, a sketch which very closely approximates the actual transient response can be constructed. A means of rapidly determining

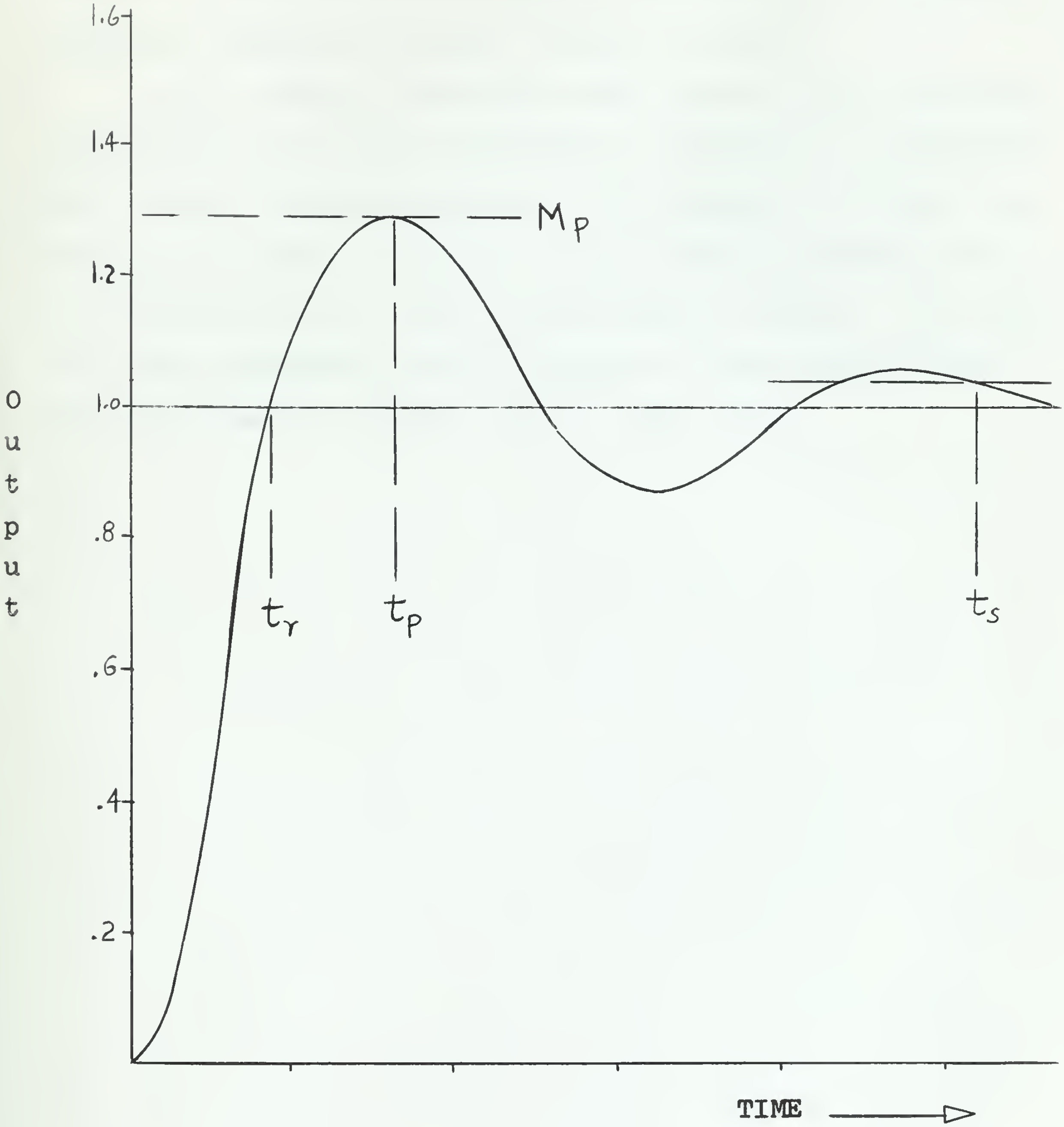

FIGURE I-1

TYPICAL TRANSIENT RESPONSE

these parameters would then be a valuable aid to the designer in evaluating the transient performance of a system without computing the exact transient response.

In this thesis a technique which permits rapid graphical determination of these parameters by inspection is developed for certain classes of systems. The technique involves the mapping of lines of constant value of these parameters onto a complex plane for a particular class of system. The values of these parameters for all systems of that class can then be read directly from the complex plane plots.

Chapter II

BASIS OF THE MAPPING TECHNIQUE

The determination of the transient response parameters of a given system is most easily approached through the performance function, the singularities of which can be displayed on a complex plane. In this approach the Laplace transform with the associated complex function theory is used as a mathematical tool.

The mapping technique developed in this thesis is based on the closed loop performance function, the poles of which are actually the roots of the system characteristic equation. The technique as developed is applicable to systems whose closed loop performance functions contain one and only one pair of complex conjugate poles.

Now consider such a performance function consisting of one pair of complex poles and an assortment of real poles and zeros. The poles and zeros of the system can be plotted on a complex plane. If the locations of the real poles and zeros are fixed, a set of values for the transient response parameters defined in Chapter I can be determined for each possible position of the complex poles. If the values of each parameter are determined for a sufficient number of complex pole locations, the loci of constant values of each

parameter can then be drawn on the complex plane. These loci are functions of the complex pole locations but are valid only for the fixed configuration of real poles and zeros.

A simple example of the loci of constant values of a parameter exists in the case of a pure second order system, that is, a system for which the closed loop performance function consists of one pair of complex poles and no real poles or zeros. In this case the overshoot is a function of damping ratio only and the lines of constant damping ratio are also lines of constant values of overshoot. These are radial lines emanating from the origin of the complex plane. This relationship is often displayed as a plot of overshoot versus damping ratio as is shown by Thaler (3).

Clearly each configuration of real poles and zeros results in a different set of loci. For example, the addition of a real pole or zero to the second order case just considered results in lines of constant overshoot which are not straight lines. In fact, the exact location and shape of the lines of constant overshoot are determined by the actual locations of the real pole or zero.

This leads to the need for a method of classifying systems according to the real pole and zero configuration. In order to facilitate the investigation of the sets of loci associated with different systems, the systems have been classified as follows:

a. Pure Second Order System - Closed loop performance function consists of two complex poles only.

b. Second Order System With One Zero - Closed loop performance function consists of one pair of complex poles and one real zero.

c. Pure Third Order System - Closed loop performance function consists of one pair of complex poles and one real pole.

These three classes of systems have been nondimensionalized so that one set of loci of constant values of parameters is sufficient for all systems of that particular class. The details of the nondimensionalizing process and the complex plane plots of loci of constant parameter values are shown in Chapters III, IV, and V. Use of the families of loci determined in these chapters requires that the closed loop performance function of a system be known in factored form.

Only the second quadrant of the complex plane is shown for two reasons.

a. Only stable systems are considered so the poles are constrained to be in the second and third quadrants.

b. The complex poles are conjugates so the third quadrant loci are images of the second quadrant loci.

In Chapter VI the technique is extended to third order systems with one or two zeros by means of an algebraic manipulation which permits evaluating the parameters of two or more systems of the classes listed above and combining the results.

Chapter III

THE PURE SECOND ORDER SYSTEM

The second order system is the simplest type to which the previously outlined mapping technique is applicable. Its characteristics have been extensively studied and almost every servomechanism text contains a complete description of its transient response. The details of this transient response will not be repeated here.

In this chapter we are concerned with what may be called a pure second order system, that is, one whose closed loop performance function consists of only two complex conjugate poles. The performance function of such a system is

$$PF(s) = \frac{C(s)}{R(s)} = \frac{a^2 + b^2}{(s+a)^2 + b^2} \qquad (1)$$

The poles of this system are shown on the s-plane in Figure III-1.

When R(t) is the unit step function the transient response of the system becomes

$$C(t) = 1 + \frac{\sqrt{a^2 + b^2}}{b} e^{-at} \sin(bt-\theta) \qquad (2)$$

where $\theta = \arctan \frac{b}{-a}$

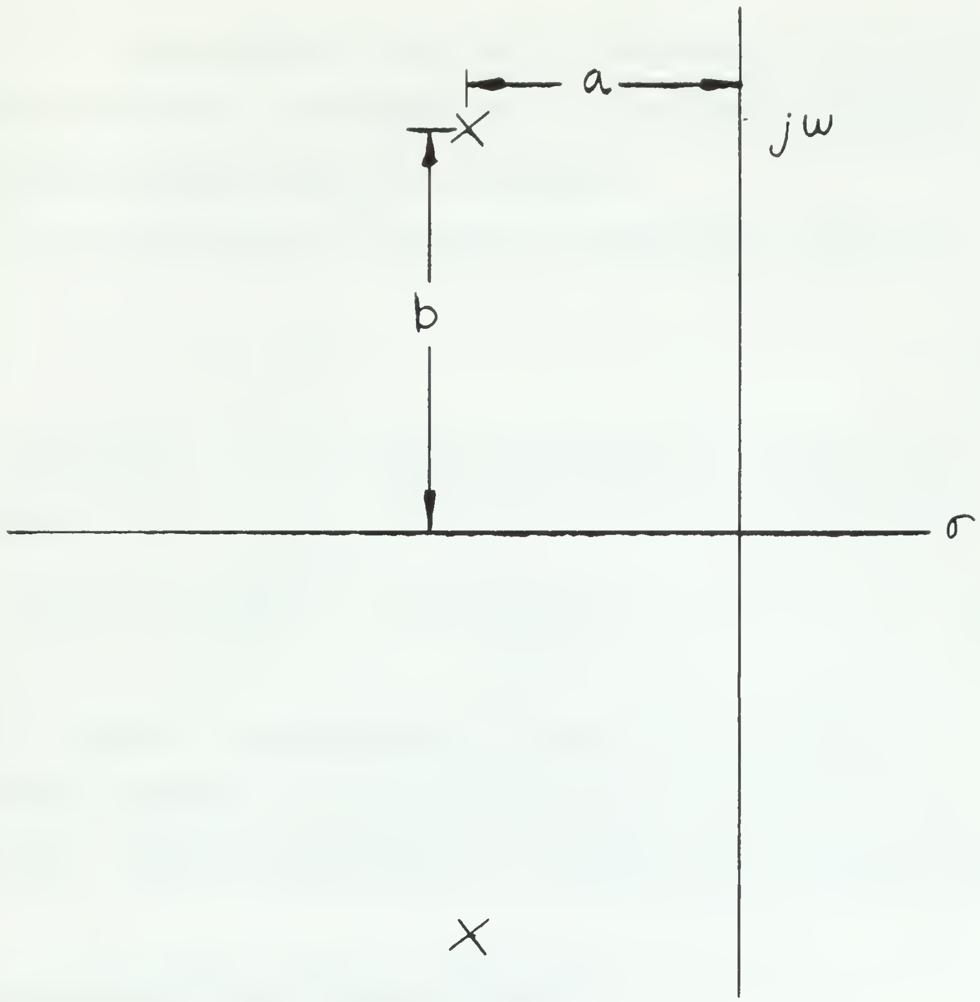

FIGURE III-1

POLES OF THE PURE SECOND ORDER SYSTEM ON THE s-PLANE

This is a very familiar expression, although it is usually written in polar coordinates, i.e., in terms of damping ratio and undamped natural frequency.

It is convenient to define an artificial time unit as

$$T = |a|\, t \tag{3}$$

The expression for the transient response in terms of T becomes

$$C(T) = 1 + \sqrt{\frac{a^2}{b^2} + 1}\;\; e^{-T} \sin\left(\frac{b}{a}T - \theta\right) \tag{4}$$

In terms of the artificial time, T, the entire transient response is determined if the ratio of a to b is known. This is equivalent to knowing the damping ratio.

<u>Time and Magnitude of Peak Overshoot</u>

The slope of the transient response is

$$\dot{C}(t) = \frac{a^2 + b^2}{b}\, e^{-at} \sin bt \tag{5}$$

or

$$\dot{C}(T) = \frac{a^2 + b^2}{b}\, e^{-T} \sin \frac{b}{a}T \tag{6}$$

From the nature of the transient response it can be seen that the peak overshoot occurs at the first time after time zero that the slope is zero. This requires that $\sin \frac{bT}{a}$ be equal to zero, or

$$\frac{b}{a}T = \pi \tag{7}$$

Therefore $$T_p = \pi\frac{a}{b} \tag{8}$$

In real time

$$t_p = \frac{T_p}{|a|} = \frac{\pi}{b} \tag{9}$$

as is shown by Truxal (4).

The loci of constant values of time of peak overshoot (T_p) are shown mapped onto the s-plane in Figure III-2. Although a scale is shown on this plot for convenience, the loci are independent of scale as they depend only on the ratio of a to b. Hence any convenient scale may be substituted for the scale shown.

The magnitude of peak overshoot can be determined by substituting T_p into the expression for C(T), Equation 4. The result of this substitution is

$$C(T_p) = 1 + e^{-\frac{\pi a}{b}} \tag{10}$$

The magnitude of peak overshoot is the value of the exponential term and once more dependent only on the ratio $\frac{a}{b}$.

$$M_p = 100\ e^{-\frac{\pi a}{b}}\ \% \tag{11}$$

Loci of constant magnitude of peak overshoot are shown mapped onto the s-plane in Figure III-3. They display the

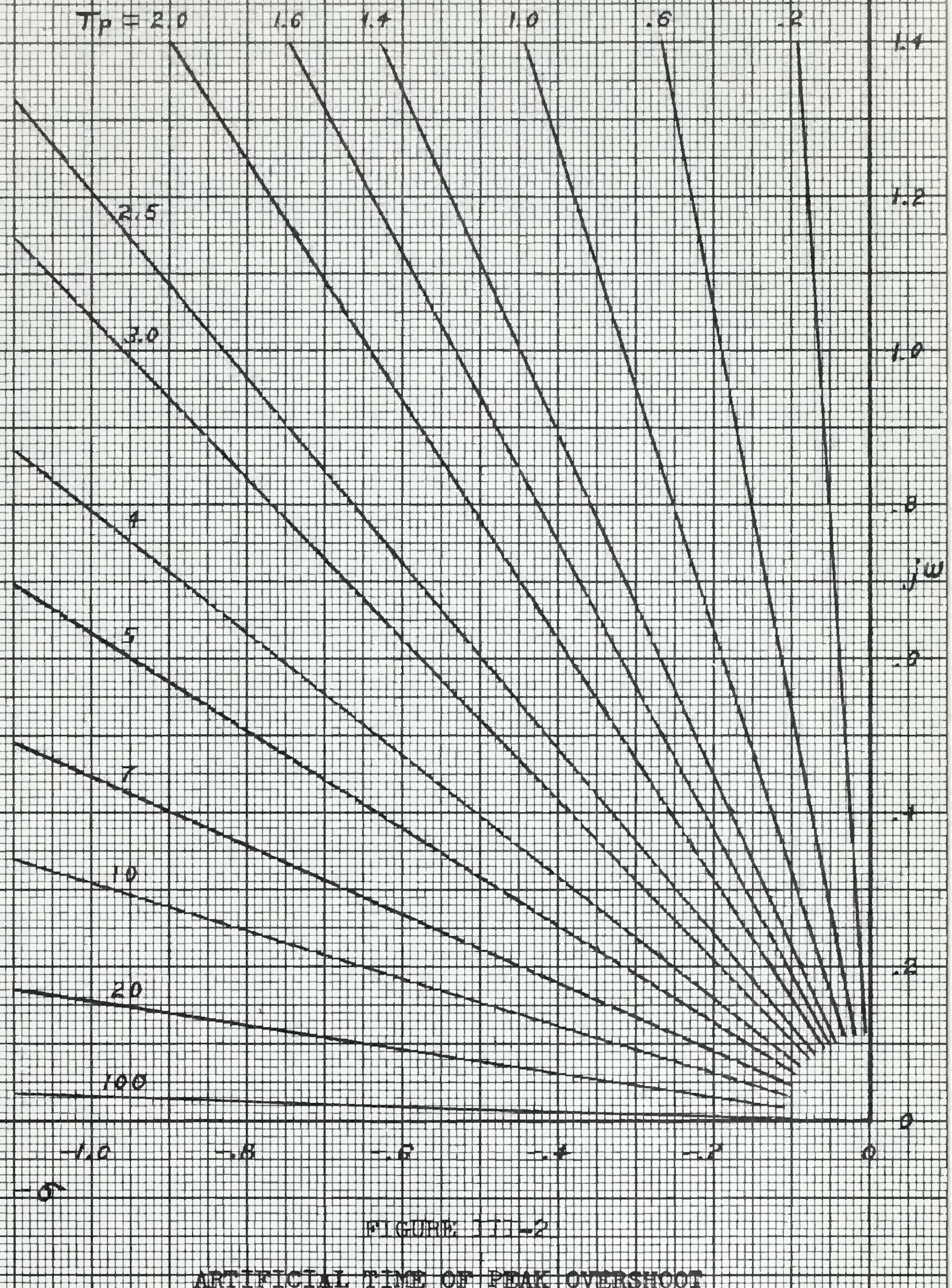

FIGURE III-2

ARTIFICIAL TIME OF PEAK OVERSHOOT

PURE SECOND ORDER SYSTEM

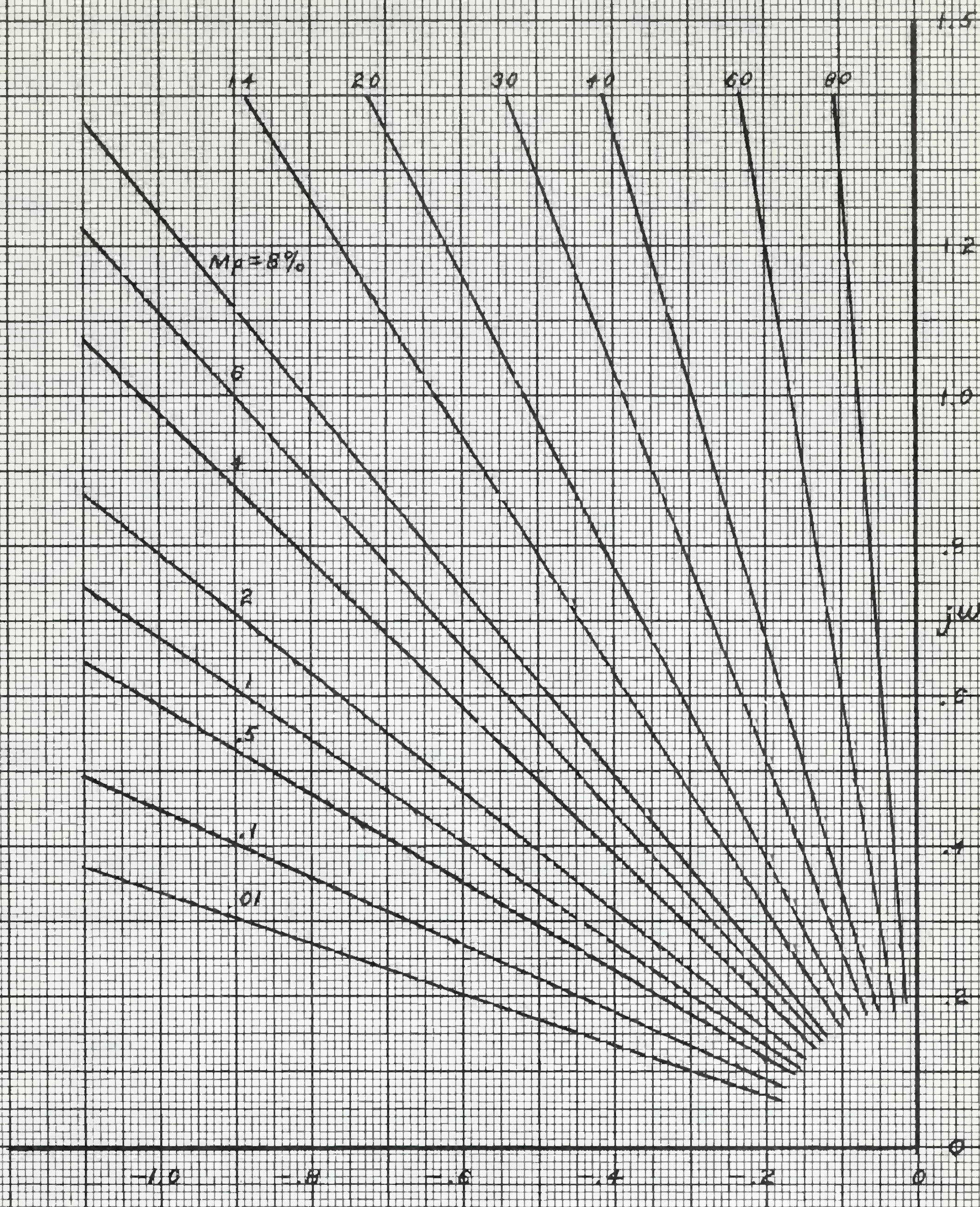

FIGURE III-3

MAGNITUDE OF PEAK OVERSHOOT - PER CENT

PURE SECOND ORDER SYSTEM

same information that is usually presented in the form of overshoot versus damping ratio plots, as has been mentioned.

Rise Time

Rise time has been defined as the first time after zero time when the value of the transient response is unity. Inspection of Equation 4 shows that this occurs when

$$\sin\left(\frac{b}{a}T - \theta\right) = 0 \tag{12}$$

or

$$\frac{b}{a}T - \theta = 0 \tag{13}$$

Therefore $$T_r = \frac{a}{b}\theta = \frac{a}{b} \arctan \frac{b}{-a} \tag{14}$$

Loci of constant values of artificial rise time (T_r) are shown mapped onto the s-plane in Figure III-4.

Settling Time

Settling time has been defined as the time when the height of the transient envelope becomes two percent of the steady state output value. This occurs when

$$\sqrt{\frac{a^2}{b^2} + 1}\; e^{-T} = .02 \tag{15}$$

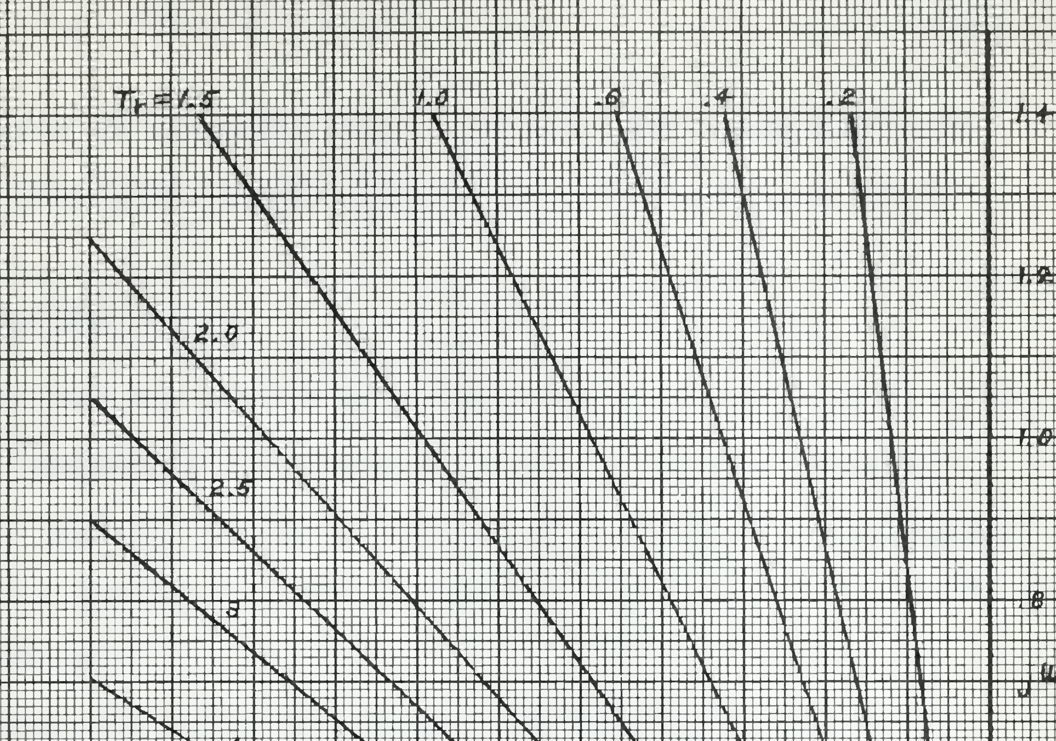

FIGURE III-4

ARTIFICIAL RISE TIME

PURE SECOND ORDER SYSTEM

Therefore

$$T_s = -\ln .02 + \frac{1}{2}\ln\left(\frac{a^2}{b^2} + 1\right) \qquad (16)$$

Loci of constant values of artificial settling time (T_s) are shown mapped onto the s-plane in Figure III-5.

<u>Use of the Curves</u>

To use the curves shown in Figures III-2 through III-5, it is necessary to determine a and b. Effectively this means that the roots of the system characteristic equation must be known. The method used in obtaining these roots is immaterial.

Once the roots are known some convenient arbitrary scale can be chosen for the loci plots. Since these are complex plane plots the real and imaginary scales must be identical. The point $-a + jb$ can then be located on each loci plot and the values of M_p, T_p, T_r, and T_s, read off directly.

The three artificial time parameters can then be converted to real times by dividing each by the value of a. The transient response may then be rapidly sketched by marking off known points, remembering that:

a. The slope is zero at $t = 0$ and $t = t_p$.

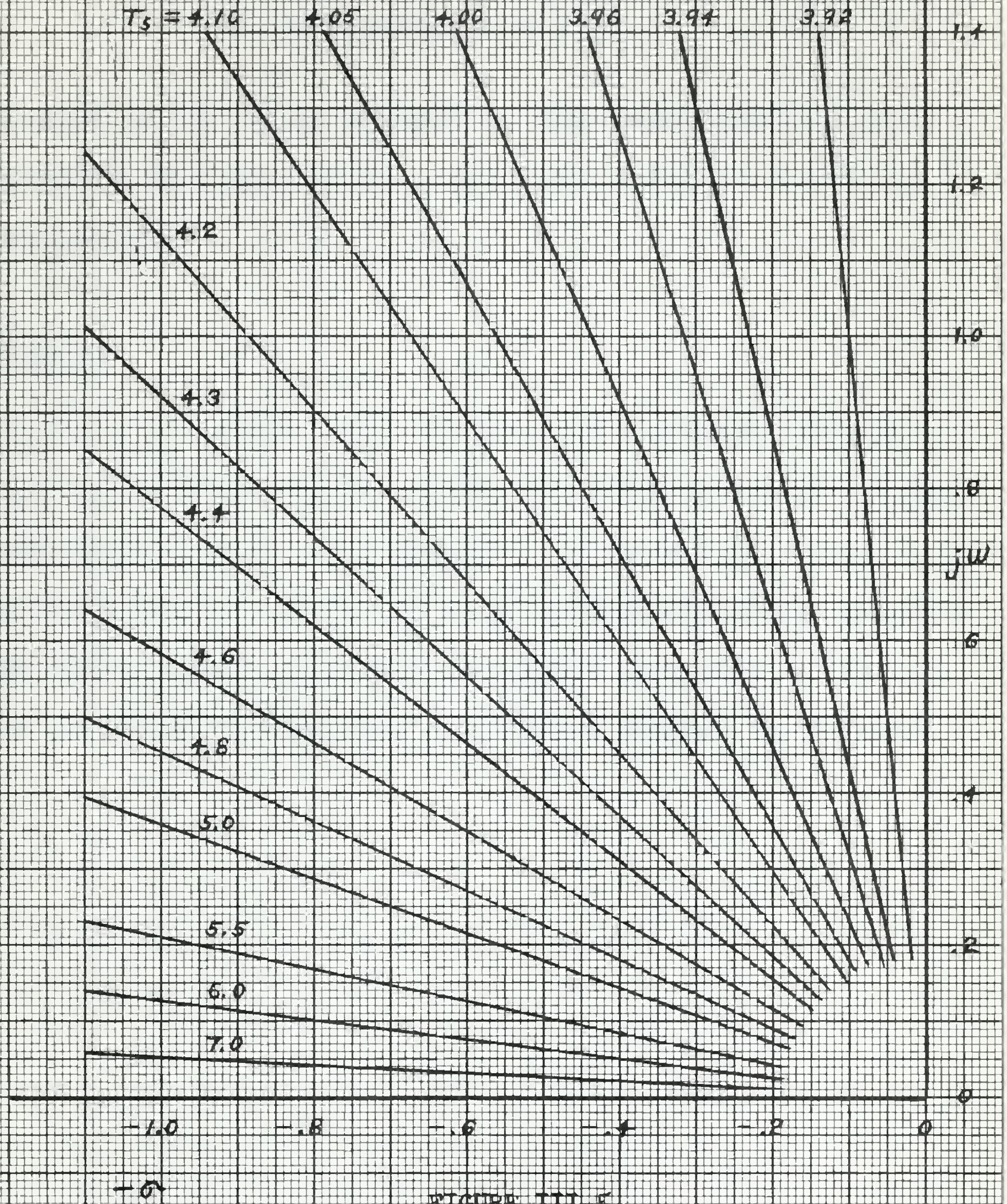

FIGURE III-5

ARTIFICIAL SETTLING TIME

PURE SECOND ORDER SYSTEM

b. The transient oscillating frequency is $\frac{b}{2\pi}$ cycles per second.

c. The sine function is zero at $t = t_r$ and increasing with time.

Figure III-6 is an example of the procedure for the performance function shown. An exact curve is included for comparison.

Conclusion

It has been shown that the loci of constant values of T_p, T_r, T_s, and M_p can be mapped directly onto the s-plane for the pure second order system. In addition, these four parameters and the transient oscillating frequency present enough information about the transient response of such a system to permit a reasonably accurate sketch to be made.

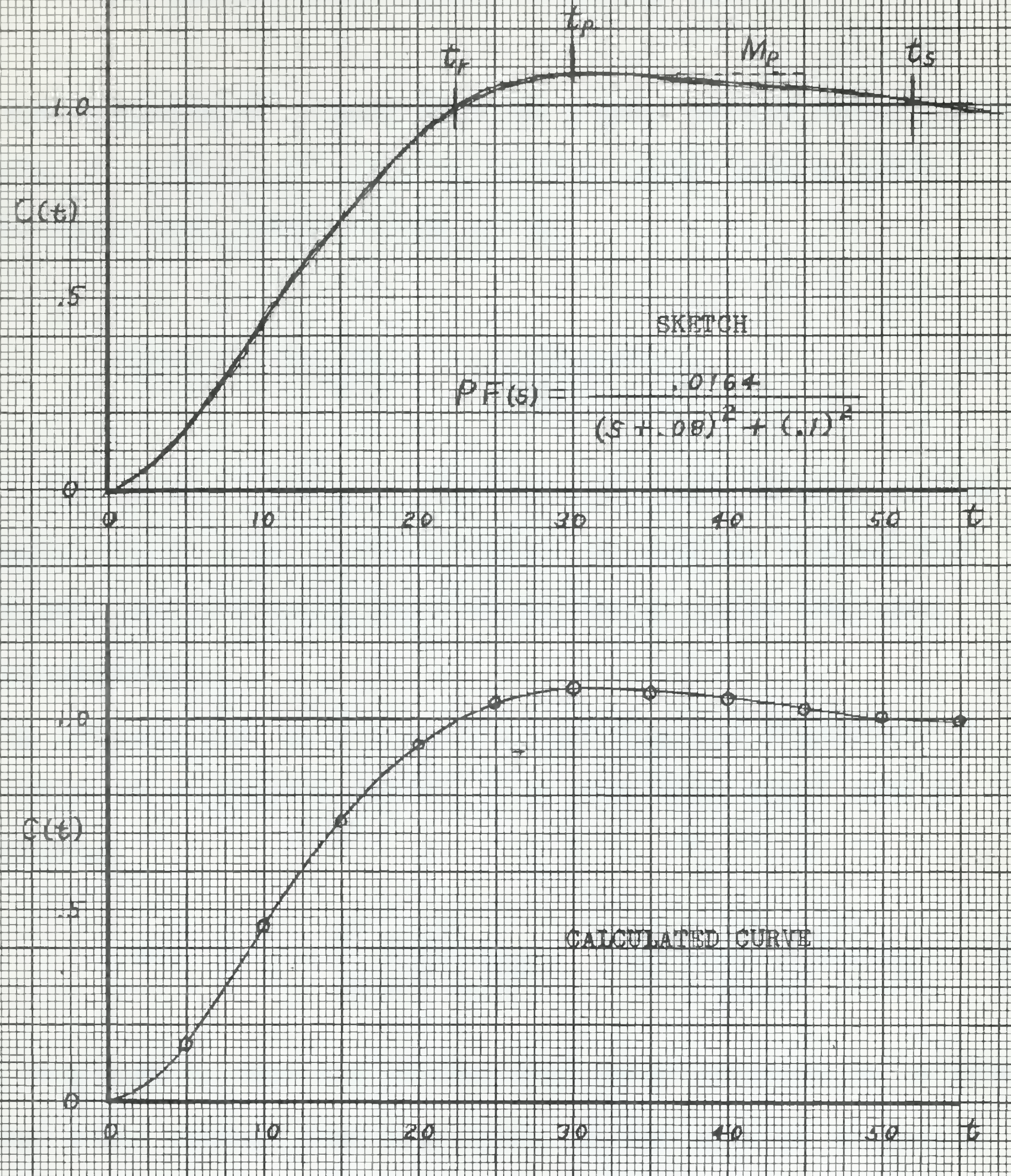

FIGURE III-6

EXAMPLE TRANSIENT RESPONSE

PURE SECOND ORDER SYSTEM

Chapter IV

THE SECOND ORDER SYSTEM WITH ONE ZERO

The addition of a real zero to a pure second order system results in a different type of second order system whose closed loop performance function is

$$PF(s) = \frac{a^2 + b^2}{d} \cdot \frac{(s + d)}{(s+a)^2 + b^2} \qquad (1)$$

The singularities of this function are shown on the s-plane in Figure IV-1. The constant which is usually interpreted as system gain reflects the assumption of unity feedback characteristics.

As pointed out by Truxal (4), the Laplace transform of the response of such a system to a unit step input may be written as

$$C(s) = \frac{1}{s} \frac{a^2 + b^2}{(s+a)^2 + b^2} + \frac{1}{d} \frac{a^2 + b^2}{(s+a)^2 + b^2} \qquad (2)$$

This response is the sum of the unit step response of a pure second order system plus a portion of the impulse response of the same system. The addition of the impulse response makes the system relatively more oscillatory than would be the case without the zero. The magnitude of this effect

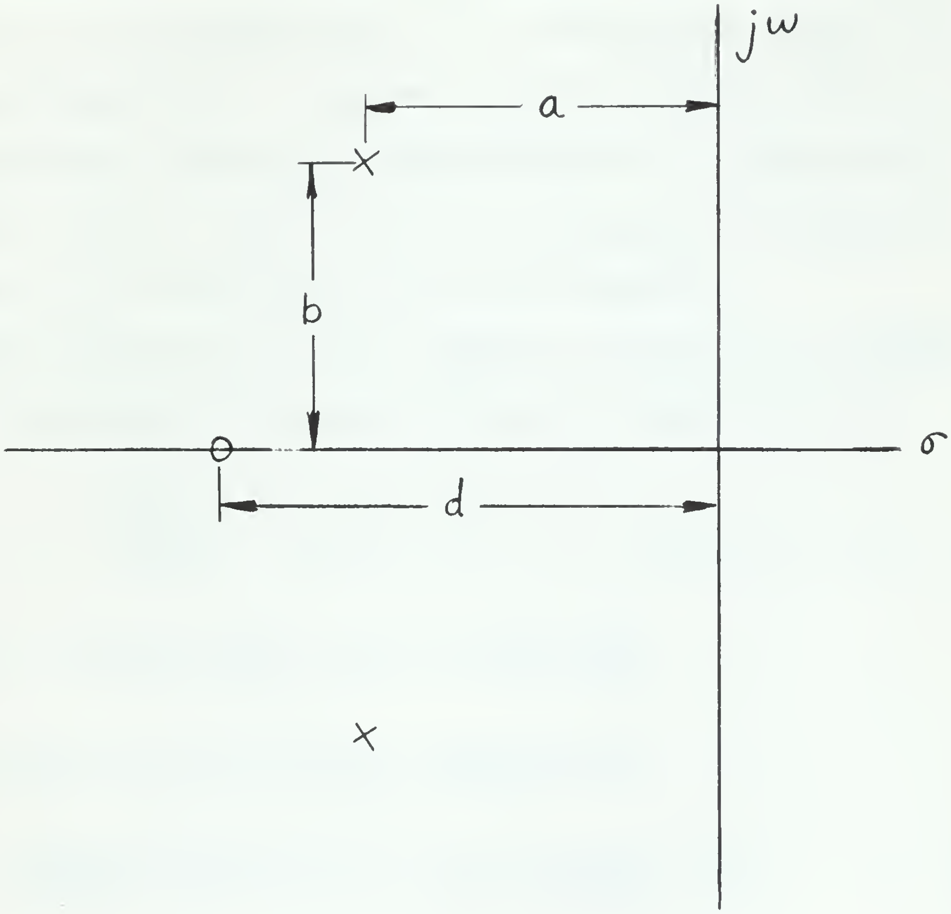

FIGURE IV-1

POLES AND ZERO ON THE s-PLANE

Second Order System with One Zero

depends on the position of the zero relative to the position of the complex poles.

It is also useful to regard the addition of the zero as a consequence of the use of proportional plus derivative control in the forward path of a conventional unity feedback servomechanism. Another equally plausible view is to consider this case as an approximation of a higher order system, that is, one in which the singularities that are far out in the left half-plane are neglected.

The expression for the transient response of a system whose performance function is given by Equation 1 is

$$C(t) = 1 + \frac{\sqrt{a^2 + b^2}}{bd} \sqrt{(d-a)^2 + b^2}\ e^{-at} \sin(bt + \phi - \theta) \qquad (3)$$

where $\phi = \arctan \frac{b}{d-a}$, and $\theta = \arctan \frac{b}{-a}$

The slope of this transient response is

$$\dot{C}(t) = \frac{a^2 + b^2}{bd} \sqrt{(d-a)^2 + b^2}\ e^{-at} \sin(bt + \phi) \qquad (4)$$

We may now express a and b in terms of d by defining

$$m_1 = \frac{a}{|d|} \qquad (5)$$

$$m_2 = \frac{b}{|d|} \qquad (6)$$

Figure IV-2 illustrates these quantities.

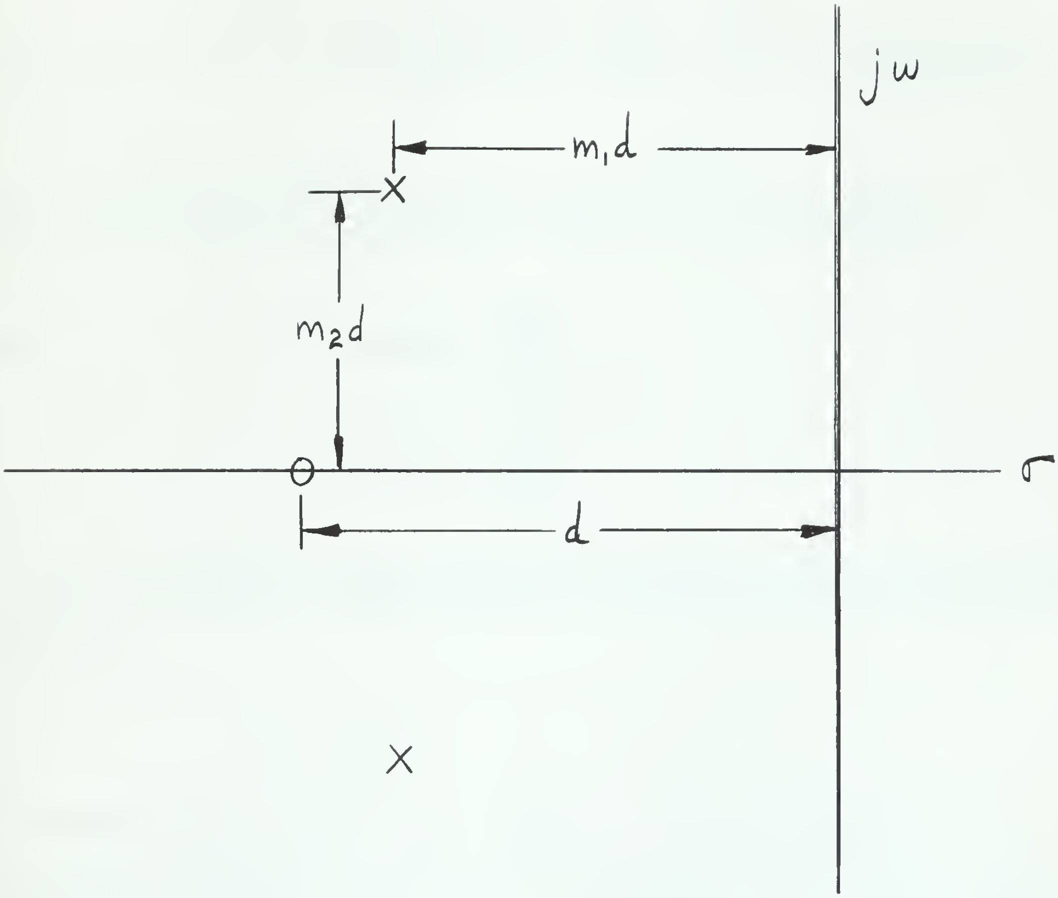

FIGURE IV-2

ILLUSTRATION OF THE QUANTITIES m_1 AND m_2

Second Order System with One Zero

Substituting these relationships into the expressions for transient response and slope of transient response, we have

$$C(t) = 1 + \frac{\sqrt{m_1{}^2 + m_2{}^2}\,\sqrt{(1-m_1)^2 + m_2{}^2}}{m_2}\, e^{-m_1 dt} \sin(m_2 dt + \phi - \Theta) \qquad (7)$$

$$\dot{C}(t) = \frac{d(m_1{}^2 + m_2{}^2)}{m_2} \sqrt{(1-m_1)^2 + m_2{}^2}\, e^{-m_1 dt} \sin(m_2 dt + \phi) \qquad (8)$$

where $\phi = \arctan \dfrac{m_2}{1-m_1}$

$\Theta = \arctan \dfrac{m_2}{-m_1}$

We now define an artificial time unit for this type of system as

$$T = |d|\, t$$

In terms of the artificial time, T, the expressions for transient response and slope of transient response become

$$C(T) = 1 + \frac{\sqrt{m_1{}^2 + m_2{}^2}\,\sqrt{(1-m_1)^2 + m_2{}^2}}{m_2}\, e^{-m_1 T} \sin(m_2 T + \phi - \Theta) \qquad (9)$$

$$\dot{C}(T) = \frac{d(m_1{}^2 + m_2{}^2)}{m_2} \sqrt{(1-m_1)^2 + m_2{}^2}\, e^{-m_1 T} \sin(m_2 T + \phi) \qquad (10)$$

where Θ and ϕ are defined as before.

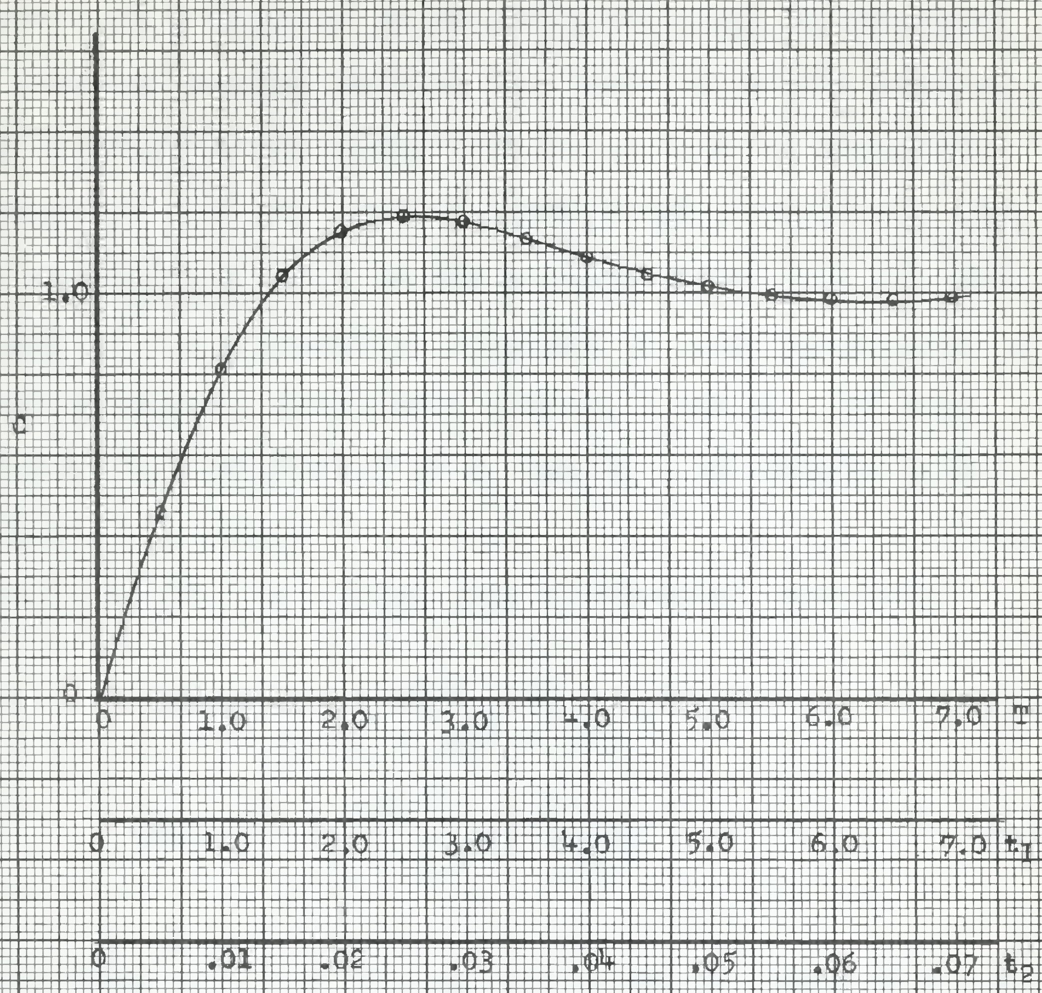

FIGURE IV-3

TRANSIENT RESPONSE OF TWO SYSTEMS WITH

SAME RELATIVE POLE-ZERO CONFIGURATION

SECOND ORDER SYSTEM WITH ONE ZERO

It is important that the reader visualize what has been done. The simplest view is that the transient response is expressed as a function of an arbitrary time, T, and that this response depends only on the position of the complex poles relative to the zero. One merely has to adopt the magnitude of d as the unit of measure and determine m_1 and m_2 by measuring a and b on the s-plane using this new unit. In terms of T the transient response is then independent of the actual numbers associated with a, b and d. All that matters is the magnitudes of m_1 and m_2. An example will illustrate the point.

Consider the two closed loop performance functions

$$PF(s)_1 = \frac{1.0(s + 1)}{(s + .6)^2 + (.8)^2} \qquad (11)$$

$$PF(s)_2 = \frac{100(s + 100)}{(s + 60)^2 + (80)^2} \qquad (12)$$

In both cases $m_1 = 0.6$ and $m_2 = 0.8$, and in terms of T the responses are identical.

$$C(T) = 1 + 1.118\, e^{-.6T} \sin(.8T + 63.4° - 126.85°) \qquad (13)$$

This response is shown in Figure IV-3 with three scales on the abcissa. The first scale is in units of T. The second is in real time and is applicable to $PF(s)_1$, where T = 1t, or $t_1 = T$. The third scale is also in real time and is applicable to $PF(s)_2$ where T = 100t or $t_2 = .01T$.

An alternate point of view is that the general performance function for this class of systems has been transformed to a new complex plane by the transformation

$$z = m_1 + jm_2 = \frac{s}{|d|} \tag{14}$$

The purpose of the transformation is to place the zero at minus one on the z-plane. All possible second order systems with one zero can be transformed in this fashion. If the variable z is treated as a Laplace variable, use of the inversion integral yields an expression for transient response, not in real time, but in an artificial time, $T = dt$. The result of this procedure is identical with the result obtained by manipulation of the expression for $C(t)$.

As has been indicated, the transient response as a function of T is dependent only on the values of m_1 and m_2. Therefore loci of constant values of magnitude of peak overshoot and the time parameters T_p, T_r, and T_s can be drawn on the z-plane.

For convenience the z-plane may be regarded as the s-plane rescaled so that the zero is placed at minus one. This places the complex conjugate poles at $-m_1 \pm jm_2$. Application of the Laplace inversion integral yields an expression in T which can then be easily converted to real time.

Magnitude and Time of Peak Overshoot

The peak overshoot occurs at the first time after zero time that the slope of the transient response is zero. This requires that

$$\sin(m_2 T + \phi) = 0 \qquad (15)$$

or

$$T_p = \frac{\pi - \phi}{m_2} = \frac{1}{m_2}\left(\pi - \arctan \frac{m_2}{1-m_1}\right) \qquad (16)$$

Loci of constant values of artificial time of peak overshoot (T_p) are shown on the z-plane in Figures IV-4 and IV-5.

To find the value of the system output at the time of peak overshoot we substitute the expression for T_p into the expression for transient response, Equation 9.

$$C(T_p) = 1 + \frac{\sqrt{m_1^2 + m_2^2}\,\sqrt{(1-m_1)^2 + m_2^2}}{m_2}\, e^{-\frac{m_1}{m_2}(\pi-\phi)} \sin(\pi - \theta) \qquad (17)$$

This reduces to

$$C(T_p) = 1 + \frac{\sqrt{m_1^2 + m_2^2}\,\sqrt{(1-m_1)^2 + m_2^2}}{m_2}\, e^{-\frac{m_1}{m_2}(\pi-\phi)} \sin\theta \qquad (18)$$

Since, as previously noted, $\tan\theta = \frac{m_2}{-m_1}$, we also know that

$$\sin\theta = \frac{m_2}{\sqrt{m_1^2 + m_2^2}} \qquad (19)$$

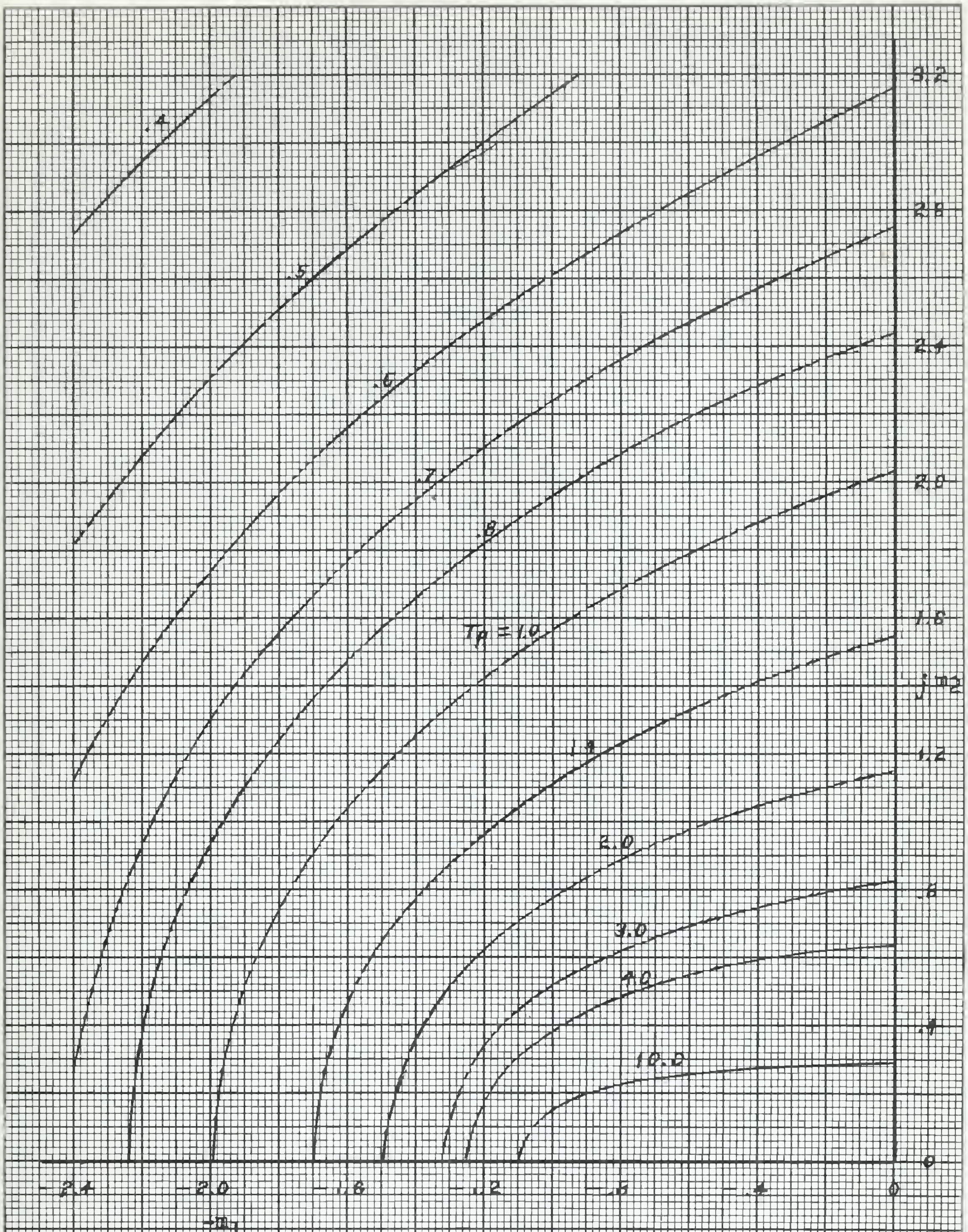

FIGURE IV-4

ARTIFICIAL TIME OF PEAK OVERSHOOT
SECOND ORDER SYSTEM WITH ONE ZERO

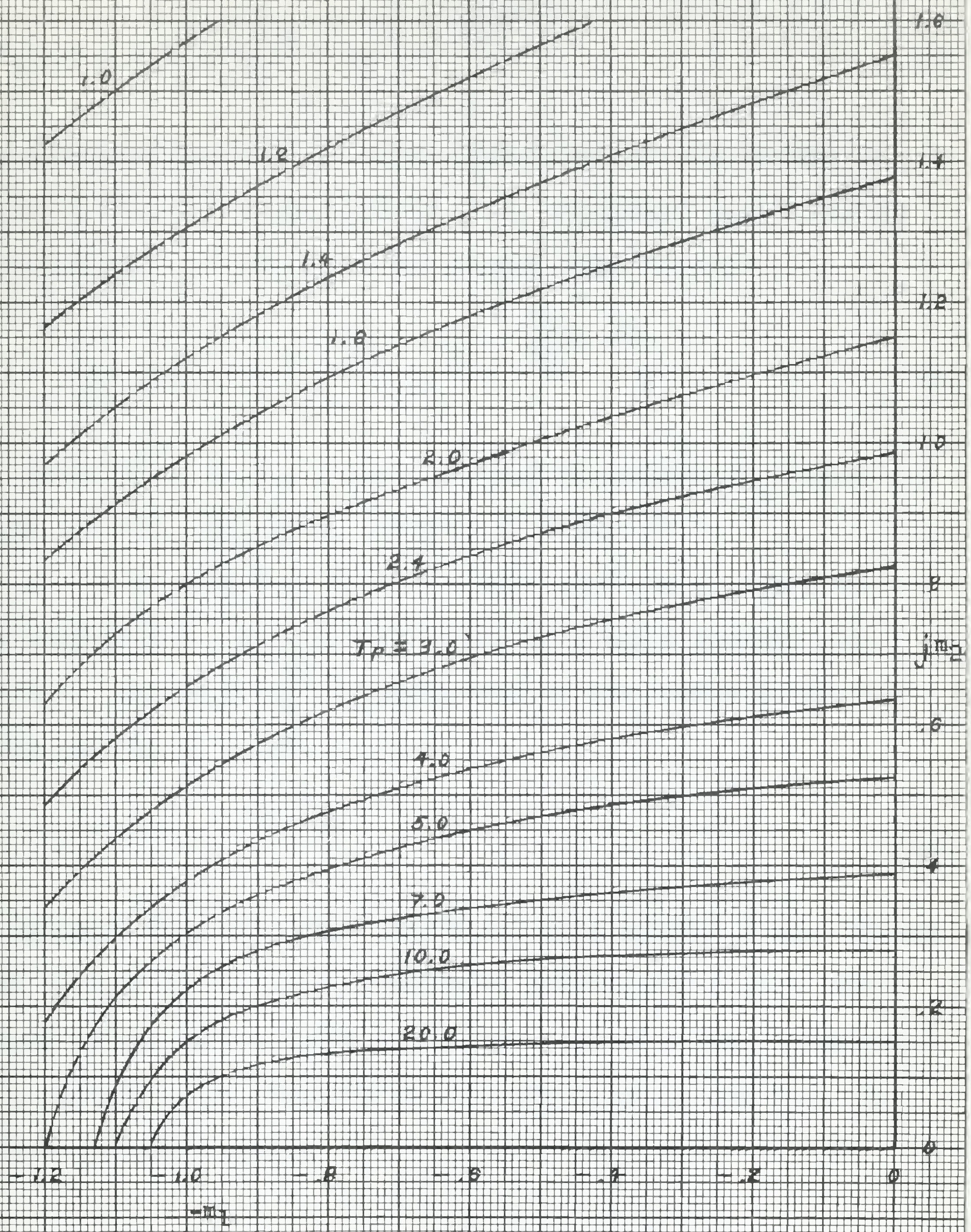

FIGURE IV-5

ARTIFICIAL TIME OF PEAK OVERSHOOT
SECOND ORDER SYSTEM WITH ONE ZERO

This can then be used to reduce Equation (18) to

$$C(T_p) = 1 + \sqrt{(1-m_1)^2 + m_2^2}\; e^{-\frac{m_1}{m_2}(\pi-\phi)} \tag{20}$$

The exponential term of this expression is the amount of peak overshoot. Magnitude of peak overshoot expressed as a percentage of steady state output then becomes

$$M_p = 100 \sqrt{(1-m_1)^2 + m_2^2}\; e^{-\frac{m_1}{m_2}(\pi-\phi)} \; \% \tag{21}$$

Loci of constant values of peak overshoot (M_p) are shown on the z-plane in Figures IV-6 and IV-7.

It is of interest to note that Equation 21 contains the expression for peak overshoot of a pure second order system. The ratio $\frac{m_1}{m_2}$ is identical to the ratio $\frac{a}{b}$ for any given values of a and b, so

$$e^{-\pi a/b} = e^{-\pi m_1/m_2} \tag{22}$$

It appears that the peak overshoot of a second order system with a zero is the peak overshoot associated with a pure second order system with the same complex roots, modified by a factor which depends on the location of the zero relative to these roots.

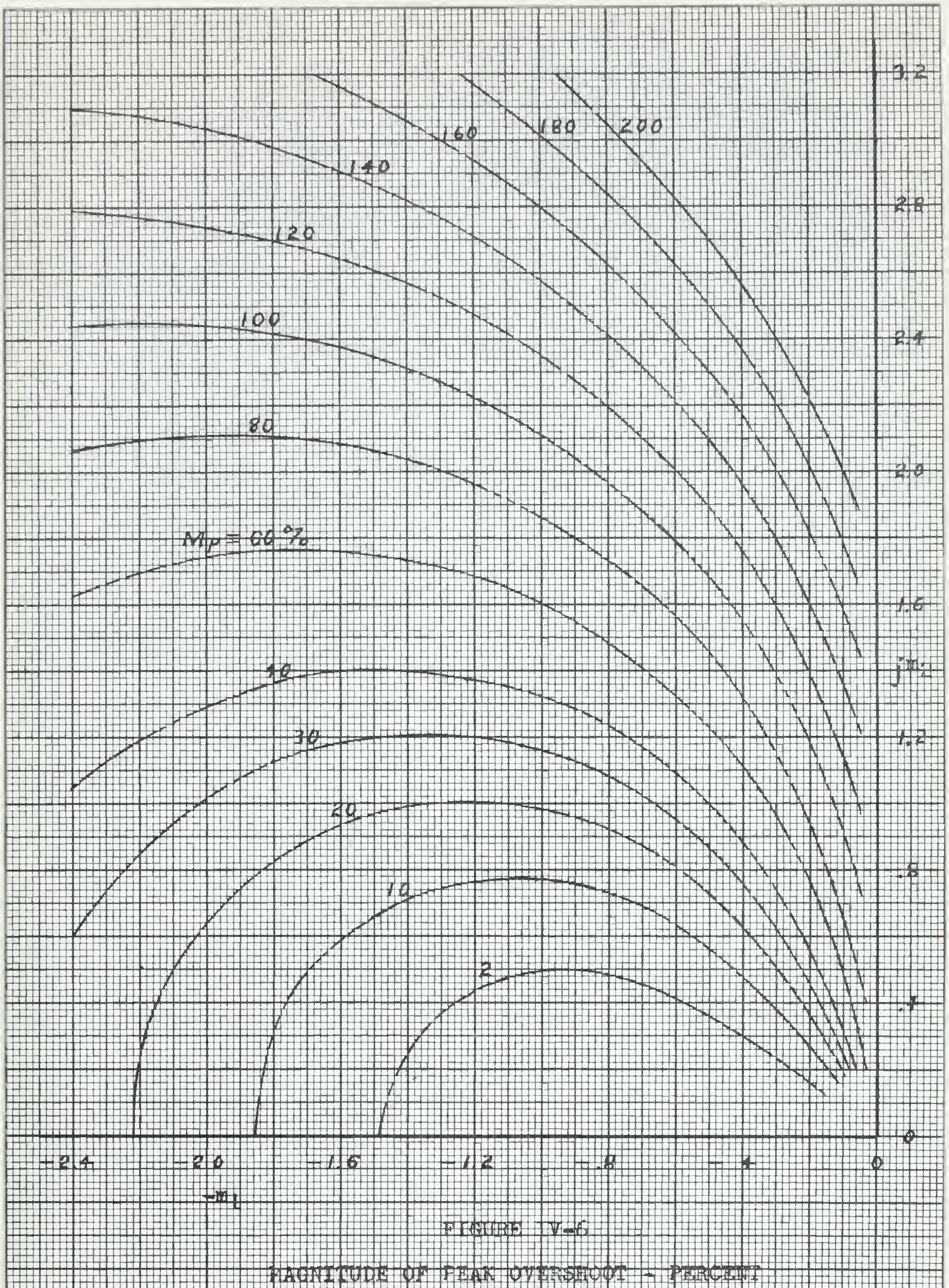

FIGURE IV-6

MAGNITUDE OF PEAK OVERSHOOT - PERCENT
SECOND ORDER SYSTEM WITH ONE ZERO

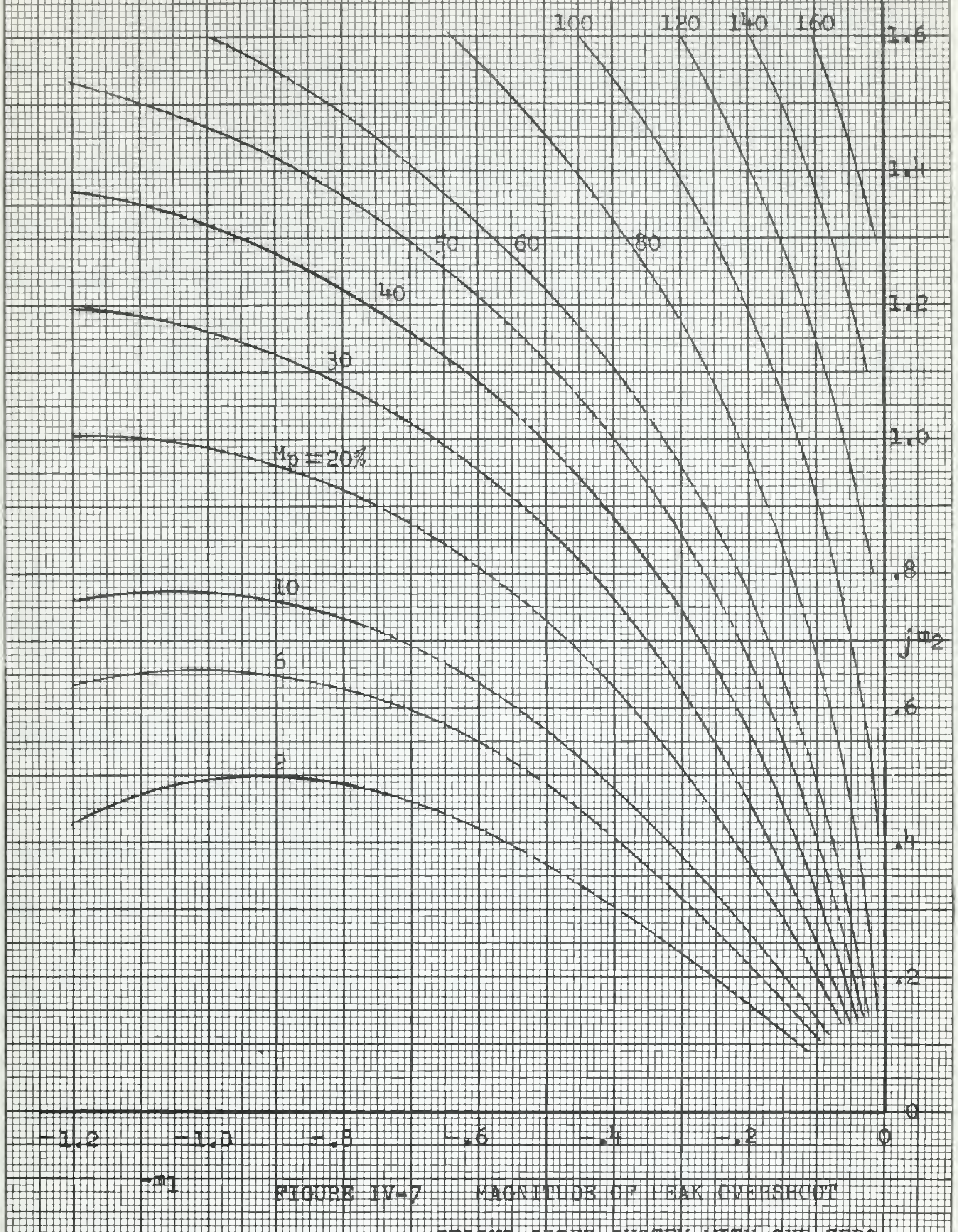

FIGURE IV-7 MAGNITUDE OF PEAK OVERSHOOT

SECOND ORDER SYSTEM WITH ONE ZERO

Rise Time

The system output is equal to its steady state output for the first time when the sine function in Equation 9 is equal to zero. The argument of the sine is always negative at zero time and increases with time. Therefore the first time after zero when the sine function is zero is when the argument is zero. That is

$$m_2T + \phi - \theta = 0 \tag{23}$$

or

$$T_r = \frac{-\phi + \theta}{m_2} \tag{24}$$

Loci of constant values of artificial rise time (T_r) are shown plotted on the z-plane in Figures IV-8 and IV-9.

Settling Time

Settling time is reached when the coefficient of the sine term in Equation 9 decays to .02 or two per cent of the value of the steady state output. This requires that

$$\frac{\sqrt{m_1^2 + m_2^2}\;\sqrt{(1-m_1)^2 + m_2^2}}{m_2}\, e^{-m_1T} = .02 \tag{25}$$

A simpler expression for T_s results if it is noted that $C(0) = 0$, as demanded by the initial conditions. In this case Equation 9 becomes

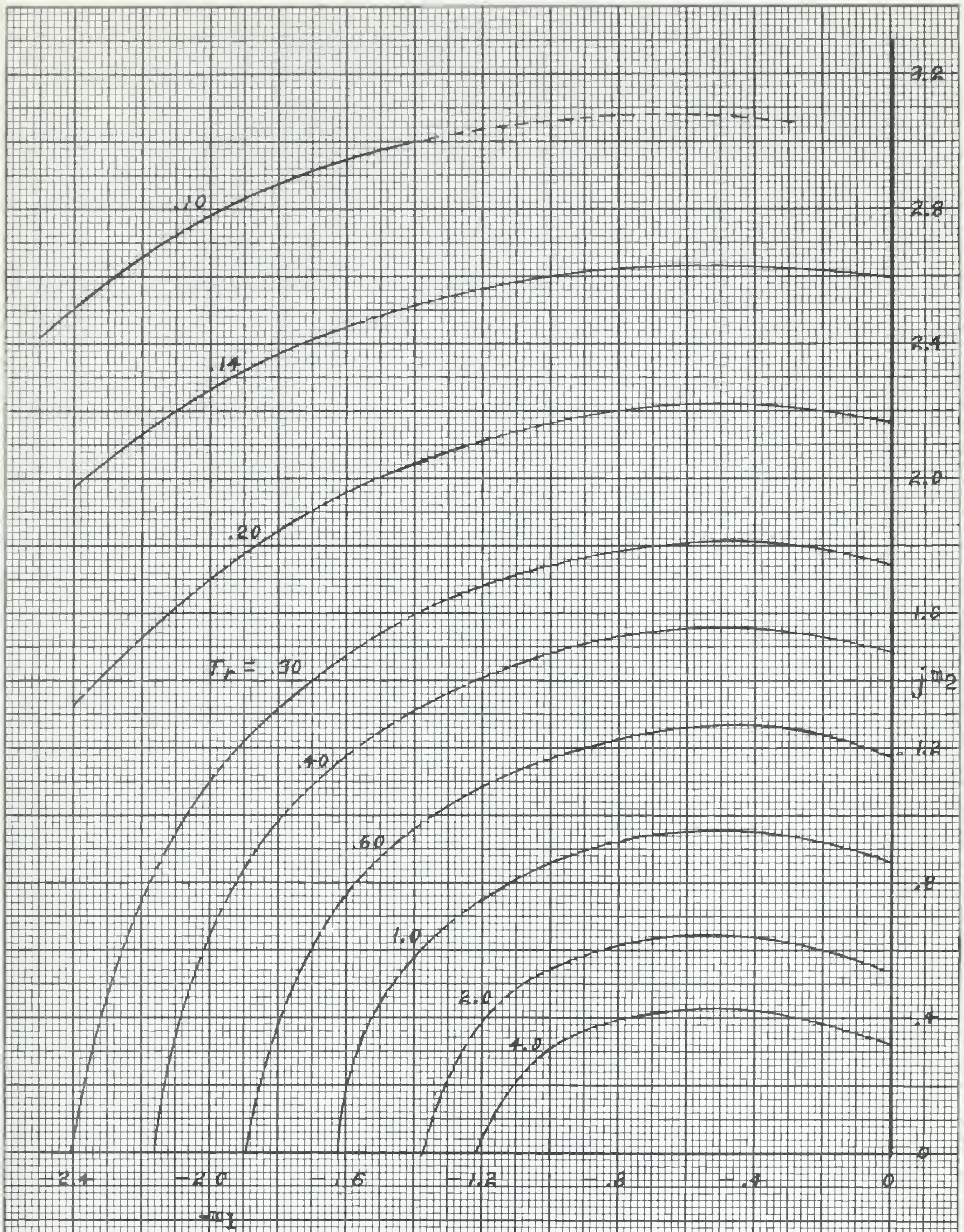

FIGURE IV-8

ARTIFICIAL RISE TIME
SECOND ORDER SYSTEM WITH ONE ZERO

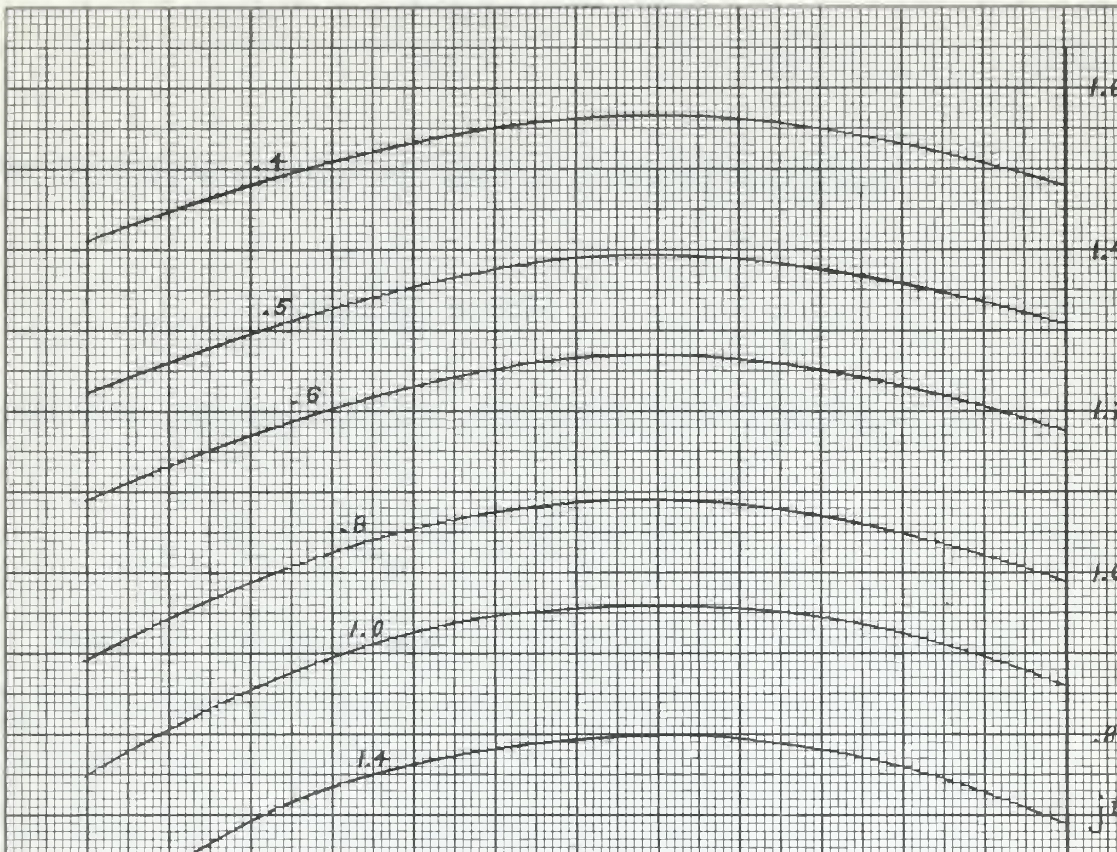

FIGURE IV-9

ARTIFICIAL RISE TIME
SECOND ORDER SYSTEM WITH ONE ZERO

$$0 = 1 + \frac{\sqrt{m_1^{\;2} + m_2^{\;2}}\sqrt{(1-m_1)^2 + m_2^{\;2}}}{m_2} \sin(\phi - \theta) \qquad (26)$$

Now Equations 25 and 26 may be combined to yield

$$- \frac{e^{-m_1 T}}{\sin(\phi-\theta)} = .02 \qquad (27)$$

From this the expression for artificial settling time is

$$T_s = - \frac{1}{m_1} \ln \left[-.02 \sin(\phi - \theta) \right] \qquad (28)$$

Loci of constant values of artificial settling time, (T_s) are shown on the z-plane in Figures IV-10 and IV-11. The loci are omitted in the zone where peak overshoot is less than two per cent because settling time as defined here is meaningless in that area.

Transient Response Parameters Along the Z-Plane Axes

The values of the transient response parameters as either m_1 or m_2 approach zero are of some interest as limiting cases. It is useful to visualize the angles θ and ϕ because they appear so often in the expressions for the parameters. They are shown in Figure IV-12.

First, consider the parameters as m_1 approaches zero, that is, along the m_2 axis or, alternately, on the s-plane along the jw axis. In the limit we have

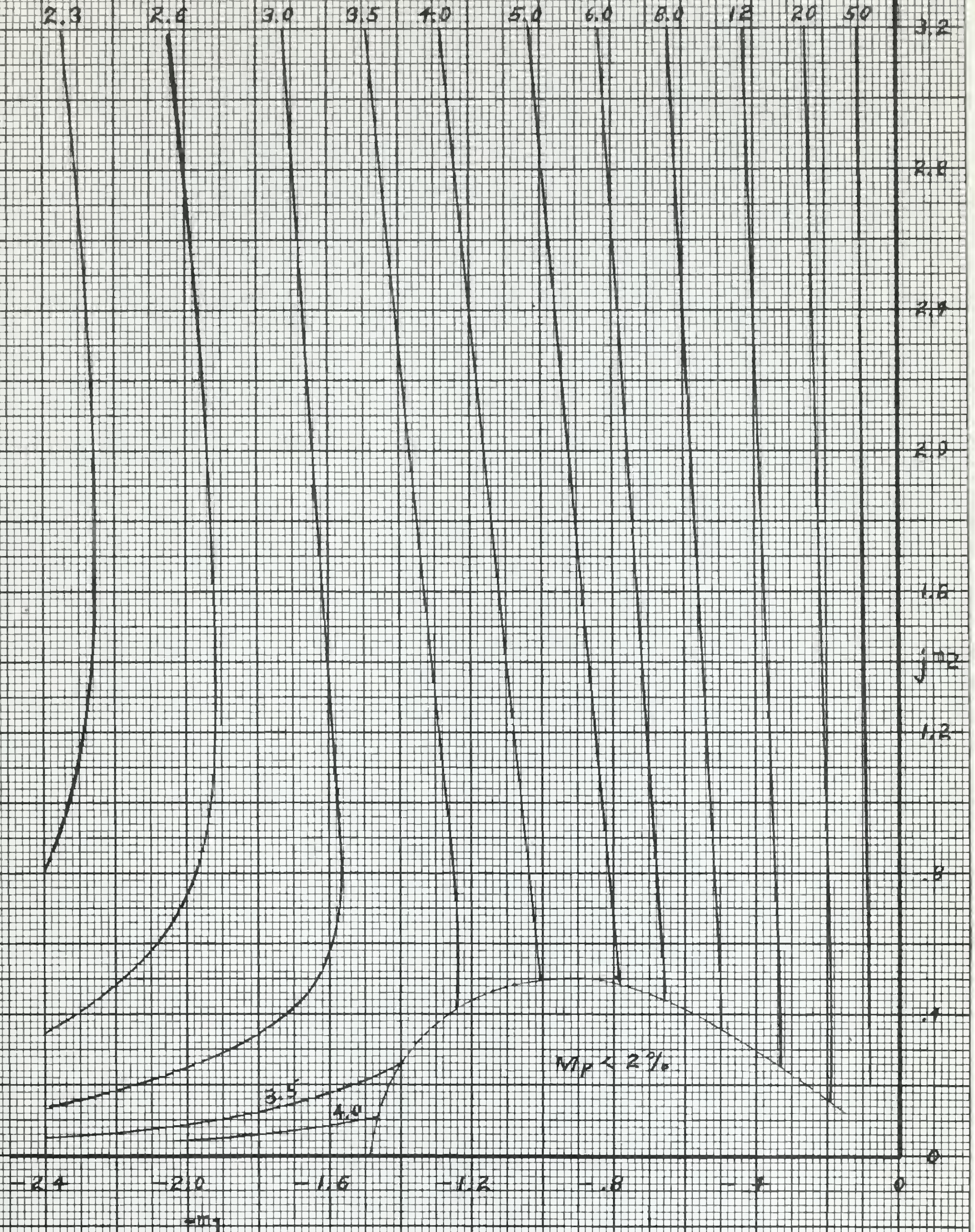

FIGURE IV-10

ARTIFICIAL SETTLING TIME
SECOND ORDER SYSTEM WITH ONE ZERO

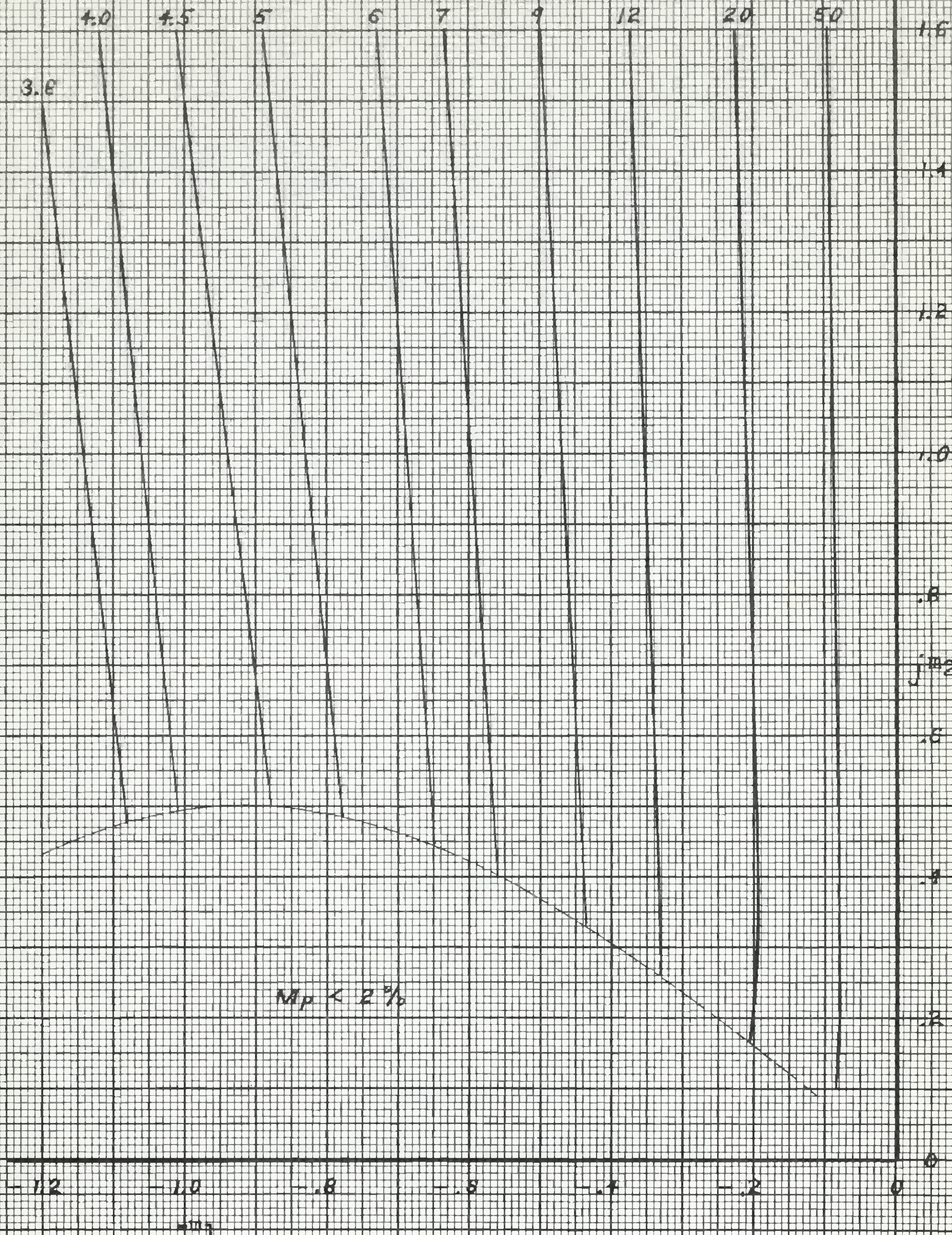

FIGURE IV-11

ARTIFICIAL SETTLING TIME
SECOND ORDER SYSTEM WITH ONE ZERO

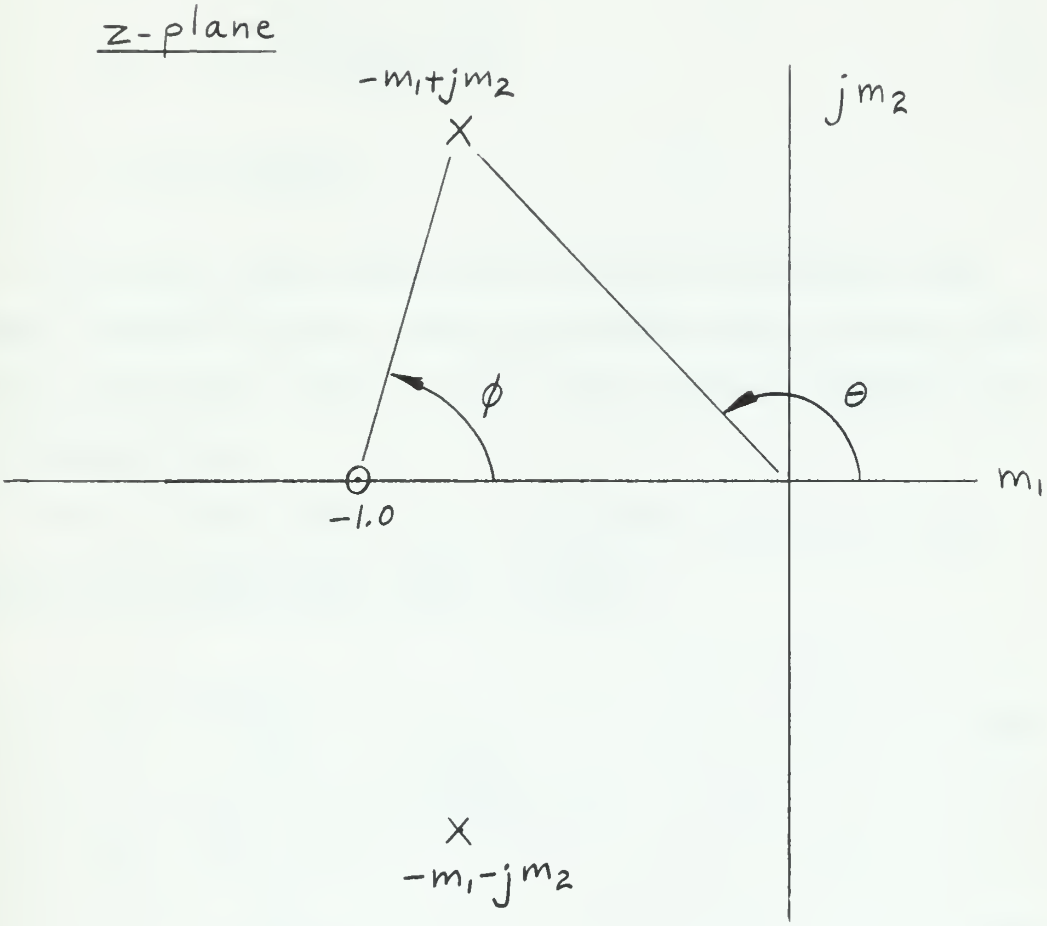

FIGURE IV-12

ANGLES Θ AND Ø ON THE z-PLANE

$$T_p = \frac{1}{m_2}(\pi - \arctan m_2) \tag{29}$$

$$M_p = 100\sqrt{1 + m_2^{\,2}}\ \% \tag{30}$$

$$T_r = \frac{-\arctan m_2 + \pi/2}{m_2} \tag{31}$$

T_s is infinite.

As m_2 approaches zero there are two distinctly different possibilities. For m_1 less than one, in the limit as m_2 approaches zero, T_p, T_r, and T_s become infinite and M_p becomes zero.

On the other hand if m_1 is greater than one, in the limit as m_2 approaches zero, we have

$$T_p = -\frac{1}{1-m_1} \tag{32}$$

$$M_p = 100(1-m_1)\,e^{+\frac{m_1}{1-m_1}}\ \% \tag{33}$$

$$T_r = -\frac{1}{m_1(1-m_1)} \tag{34}$$

T_s is infinite.

These expressions are useful in determining the intercepts of a particular locus.

Construction of the Loci

Each of the families of loci were constructed by calculating and plotting the values of the parameters as functions of m_2 for a number of fixed values of m_1. The result was a series of cross-sections of the surfaces defined by the parameters over the z-plane. The coordinates m_1 and m_2 were then read from these cross-section curves for each locus desired, and these points were plotted on the z-plane. This method was required because no method could be found to determine the equations of the loci on the z-plane.

Use of the Curves

The curves are used to determine the transient response parameters in the same manner as outlined in Chapter III in connection with pure second order systems. Once the closed loop performance function is known in factored form, that is, the quantities a, b, and d are determined, m_1 and m_2 must be calculated. This is relatively simple since

$$m_1 = \frac{a}{|d|} \quad \text{and} \quad m_2 = \frac{b}{|d|}$$

The point $-m_1 + jm_2$ is then located on each family of loci and values of T_p, T_r, T_s, and M_p are read off directly.

The three time parameters can then be converted to real time by dividing each by the value of d, i.e.,

$$t_p = \frac{T_p}{|d|}, \qquad t_r = \frac{T_r}{|d|}, \qquad \text{and} \quad t_s = \frac{T_s}{|d|}$$

With these three times and the magnitude of peak overshoot the transient response may be rapidly sketched if it is remembered that:

a. The slope at $t = t_p$ is zero.

b. The transient oscillating frequency is $\frac{b}{2\pi}$ cycles per second.

c. The sinusoid is zero at $t = t_r$, and is increasing with time.

d. The slope at zero time is $\frac{a^2+b^2}{d}$, which can be easily shown by use of the initial value theorem.

Observations

The loci of constant values of peak overshoot become straight lines as they approach the origin. If m_1 and m_2 are much less than one, the expression for peak overshoot degenerates to

$$M_p = 100 \ e^{-\frac{m_1}{m_2}\pi} \ \%$$

As has been shown, this is identical to the expression for peak overshoot for pure second order systems. These loci are actually very nearly straight lines if m_1 and m_2 are less than about 0.6, indicating that a zero will not invalidate a pure second order approximation unless its magnitude is less than about twice the system undamped natural frequency. For very low values of <u>both</u> m_1 and m_2 the loci of constant values of parameters have been omitted because the pure second order curves are very nearly exact in this region.

A comparison of the loci of constant values of peak overshoot shown in Figure IV-7 with the same loci for pure second order systems as shown in Figure III-3 will illustrate the previous observation. The comparison will also amplify the phrase "relatively more oscillatory than would be the case without the zero". For fixed complex root locations the addition of a zero can cause a large increase in the magnitude of peak overshoot. In fact, peak overshoot of greater than one hundred percent is a distinct possibility for such systems. It should be noted that the addition of a zero does not effect the transient oscillating frequency.

The initial slope of the transient response is not zero as was the case with the pure second order systems. The initial conditions are met, however, because there is a velocity discontinuity at zero time due to the combination

of the input function and the zero. As was mentioned, the zero can arise from the addition of derivative control in the forward path. Differentiation of the unit step through this channel yields a unit impulse, and it is this impulse that accounts for the initial slope.

Chapter V

THE PURE THIRD ORDER SYSTEM

The addition of a real root to the pure second order system results in a new system which we define as a pure third order system. The closed loop performance function of the pure third order system then has one pair of complex conjugate poles, one real pole and no zeros.

Much less is known about the pure third order system than is the case with the pure second order system. This is due largely to the introduction of a third variable in the equations due to the third root. The mathematical expression for the transient response of the pure third order system contains a decaying exponential term in addition to the constant and damped sinusoidal terms of the second order system. The actual expression for the transient response is not difficult to obtain, nor to plot, but rather the actual plotting is laborious. The main difficulty in establishing standard sets of curves for the third order system, as is often done with the second order system, is in portraying these curves as functions of three variables.

This does not imply that nothing has been done in this area. To the contrary, much effort has gone into the

study of third order systems. Unfortunately, the results of this effort has often produced new ways of factoring the system characteristic equations without producing any simple method of evaluating the transient response. One such method was developed by Y. J. Liu (5), and is based on nondimensionalizing the characteristic equation. By using this method, stability and the roots of the characteristic equation can be determined from a set of nondimensionalized charts.

Another method of importance in any discussion of third order systems is Mitrovic's method (6). In this method, by use of a selected transformation, it is possible to compute a set of curves which inherently contain the solution to all third order equations and can be used to determine stability. These curves have been computed and are known as third order charts. They may be used directly in compensation by cancellation provided suitable root locations are known. However, this still does not solve the problem of where to locate the roots and it does not yield transient response directly.

Transient Response

A plot of the roots of the pure third order system is shown in Figure V-1. The closed loop performance function for the unity feedback third order system may be written as

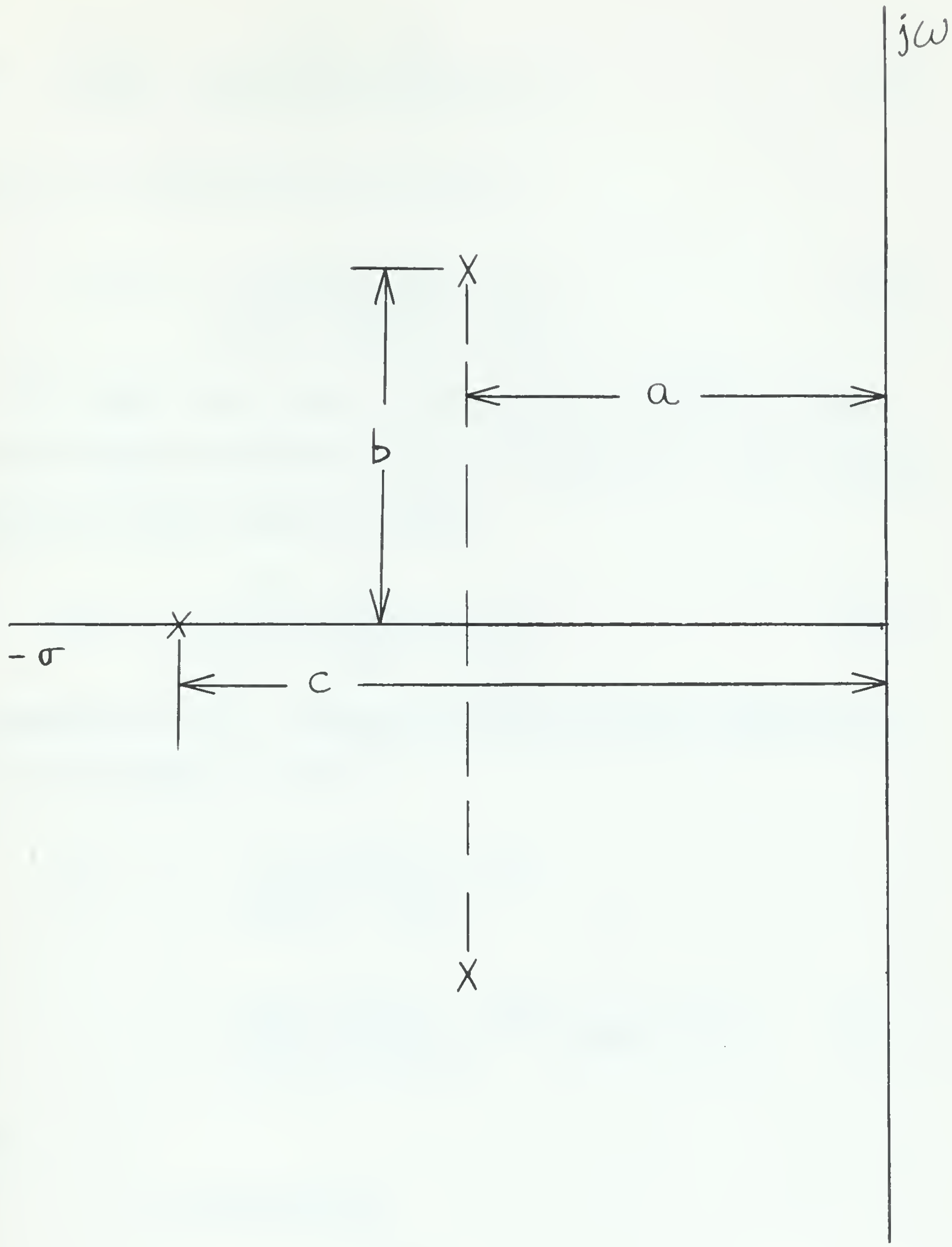

FIGURE V-1

POLES OF THE PURE THIRD

ORDER SYSTEM ON THE s-PLANE

$$PF(s) = \frac{C(s)}{R(s)} = \frac{(a^2 + b^2)c}{(s+c)(s+a+jb)(s+a-jb)} \qquad (1a)$$

This may be simplified as follows

$$PF(s) = \frac{(a^2 + b^2)c}{(s+c)\left[(s+a)^2 + b^2\right]} \qquad (1b)$$

If a unit step input is applied to a system having the performance function shown in Equation 1, the complex frequency domain output becomes

$$C(s) = \frac{(a^2 + b^2)c}{s(s+c)\left[(s+a)^2 + b^2\right]} \qquad (2)$$

The inverse of this transform gives the time domain output of the system, which is

$$C(t) = 1 - \frac{(a^2 + b^2)}{\left[(a-c)^2 + b^2\right]} e^{-ct} + \frac{c\sqrt{a^2 + b^2}}{b\sqrt{(c-a)^2 + b^2}} e^{-at} \sin(bt-\theta-\phi) \qquad (3)$$

where

$$\theta = \arctan \frac{b}{-a}$$

$$\phi = \arctan \frac{b}{c-a}$$

Now if we redefine the root locations in terms of the location of the real root as shown in Figure V-2, the real

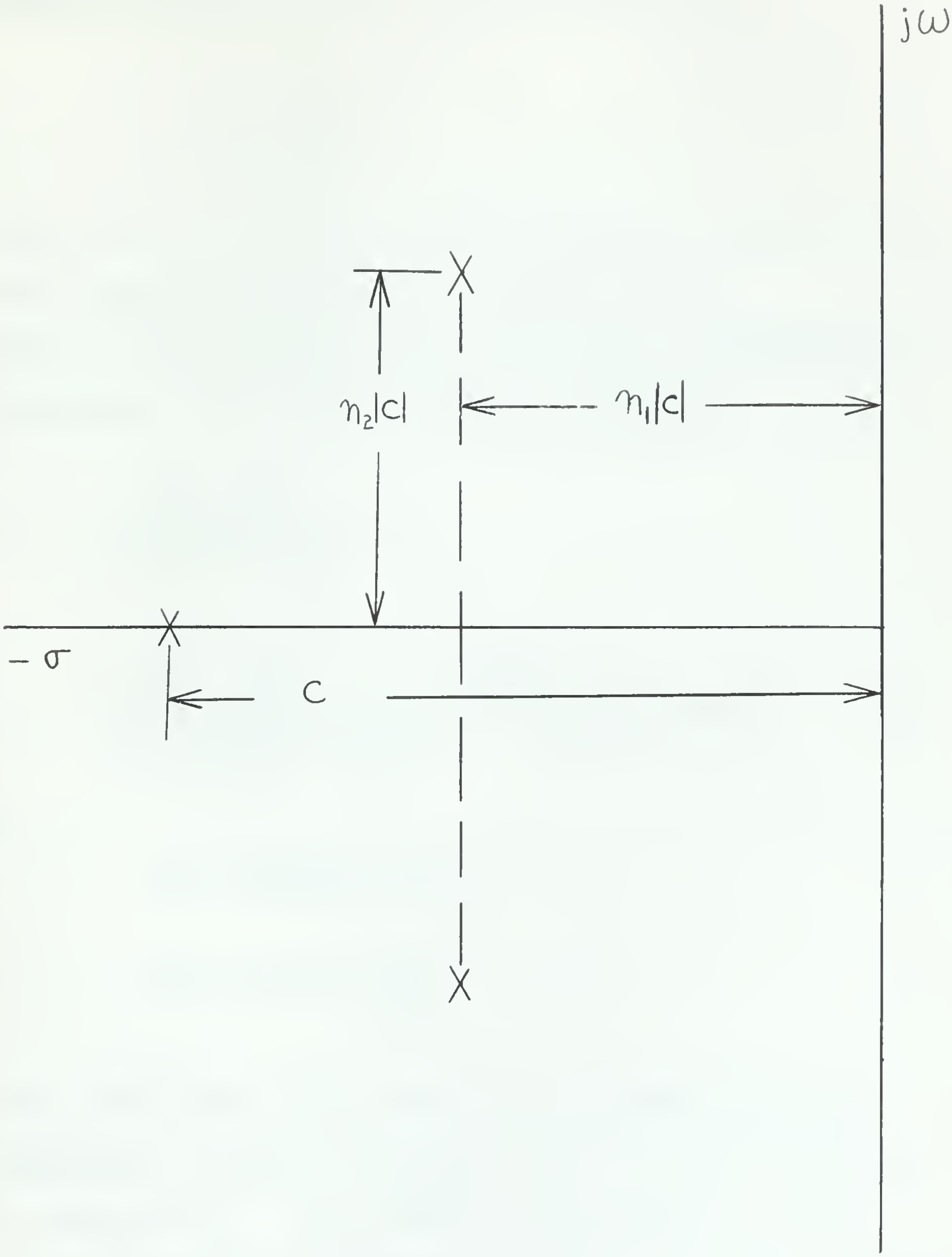

FIGURE V-2

ILLUSTRATION OF QUANTITIES n_1 AND n_2

PURE THIRD ORDER SYSTEM

and imaginary parts of the complex roots become

$$a = n_1 |c|$$

$$b = n_2 |c|$$

The quantities n_1 and n_2 are merely rescaled values of the real and imaginary parts of the complex roots.

Making these substitutions for a and b in Equation 3, the expression for transient response becomes

$$C(t) = 1 - \frac{n_1^{\ 2} + n_2^{\ 2}}{(1-n_1)^2 + n_2^{\ 2}} e^{-ct}$$

$$+ \frac{\sqrt{n_1^{\ 2} + n_2^{\ 2}}}{n_2 \sqrt{(1-n_1)^2 + n_2^{\ 2}}} e^{-n_1 ct} \sin(n_2 ct - \theta - \phi) \qquad (4)$$

where

$$\theta = \arctan n_2/-n_1$$

$$\phi = \arctan \frac{n_2}{1-n_1}$$

This substitution has partially nondimensionalized the transient response of systems with the same relative root configuration. The transient response can now be completely nondimensionalized by making the additional substitution $T = |c|t$. This is equivalent to defining a new artificial time for this class of systems as was done in Chapters III and IV. The transient response now becomes

$$C(T) = 1 - \frac{(n_1^{\,2} + n_2^{\,2})}{(1-n_1)^2 + n_2^{\,2}} e^{-T} \tag{5}$$

$$+ \frac{\sqrt{n_1^{\,2} + n_2^{\,2}}}{n_2 \sqrt{(1-n_1)^2 + n_2^{\,2}}} e^{-n_1 T} \sin(n_2 T - \theta - \phi)$$

It is of interest to look at a different derivation of Equation 5. As a first step let us write the closed loop performance function in terms of the root locations shown in Figure V-2.

$$PF(s) = \frac{(n_1^{\,2} + n_2^{\,2})c^3}{(s+c)\left[(s+n_1 c)^2 + (n_2 c)^2\right]} \tag{6a}$$

Dividing numerator and denominator by c^3 gives

$$PF(s) = \frac{n_1^{\,2} + n_2^{\,2}}{(\frac{s}{c} + 1)\left[(\frac{s}{c} + n_1)^2 + n_2^{\,2}\right]} \tag{6b}$$

Now if we make the linear transformation

$$Z = n_1 + jn_2 = \frac{s}{|c|}$$

the performance function becomes

$$PF(z) = \frac{n_1^{\,2} + n_2^{\,2}}{(z+1)\left[(z+n_1)^2 + n_2^{\,2}\right]} \tag{7}$$

This has effectively transformed the system onto a new complex plane in which the real root is located at

minus one and the complex roots are at $-n_1 \pm jn_2$. Now upon application of a unit step the complex frequency domain output becomes

$$C(z) = \frac{n_1^2 + n_2^2}{z(z+1)\left[(z+n_1)^2 + n_2^2\right]} \tag{8}$$

The inverse transform of this gives the transient response as

$$\begin{aligned} C(T) = 1 &- \frac{n_1^2 + n_2^2}{(1-n_1)^2 + n_2^2} e^{-T} \\ &+ \frac{\sqrt{n_1^2 + n_2^2}}{n_2 \sqrt{(1-n_1)^2 + n_2^2}} e^{-n_1 T} \sin(n_2 T - \theta - \emptyset) \end{aligned} \tag{9}$$

In this case T is the time associated with the inverse from the z-plane. Note that Equation 9 is identical to Equation 5.

Time and Magnitude of Peak Overshoot

As in the previous cases, the slope of the transient response is zero at the time of peak overshoot. That is

$$\frac{d}{dt} C(t_p) = 0 \tag{10}$$

The Laplace transform of this for zero initial conditions is

$$\mathcal{L}\left[\frac{d}{dt}\, C(t)\right] = S\, C(s) \tag{11}$$

Multiplying both sides of Equation 2 by s gives

$$S\, C(s) = \frac{(a^2 + b^2)c}{(s+c)\left[(s+a)^2 + b^2\right]} \tag{12}$$

Combining Equations 11 and 12, and obtaining the inverse transform gives the slope of the transient response as

$$\frac{d}{dt}\, C(t_p) = \frac{(a^2 + b^2)c}{(c-a)^2 + b^2}\, e^{-ct_p}$$

$$+ \frac{(a^2 + b^2)}{b\sqrt{(c-a)^2 + b^2}}\, e^{-at_p} \sin(bt_p - \phi) = 0 \tag{13}$$

where ϕ is defined as before $\qquad \phi = \arctan \frac{b}{c-a}$

Equation 13 can be reduced to

$$\frac{c(a^2 + b^2)\left[b\, e^{-ct_p} + \sqrt{(c-a)^2 + b^2}\, e^{-at_p} \sin(bt_p - \phi)\right]}{b\left[(c-a)^2 + b^2\right]} = 0 \tag{14}$$

For practical systems the left side of Equation 14 can equal zero only when the term in brackets equals zero. Therefore, the condition to be satisfied at time of peak overshoot is

$$b\, e^{-ct_p} + \sqrt{(c-a)^2 + b^2}\, e^{-at_p} \sin(bt_p - \phi) = 0 \tag{15}$$

Equation 15 can now be nondimensionalized by making the same substitutions as used in the transient response, i.e.,

$$a = n_1|c|$$
$$b = n_2|c|$$
$$T = |c|t$$

When these substitutions are made in Equation 15 the condition to be satisfied at time of peak overshoot becomes

$$\frac{-n_2}{\sqrt{(1-n_1)^2 + n_2^2}} e^{-(1-n_1)T} = \sin(n_2 T - \phi) \tag{16}$$

where $\phi = \arctan \frac{n_2}{1-n_1}$

From Equation 5 we see that the magnitude of peak overshoot (expressed as a percentage) can be written as

$$M_p = 100 \left[\frac{\sqrt{n_1^2 + n_2^2}}{n_2 \sqrt{(1-n_1)^2 + n_2^2}} e^{-n_1 T_p} \sin(n_2 T_p - \theta - \phi) - \frac{n_1^2 + n_2^2}{(1-n_1)^2 + n_2^2} e^{-T_p} \right] \% \tag{17}$$

For any combination of n_1 and n_2, Equation 16 can be solved by trial and error techniques to determine the values of T for the first, second and each succeeding maxi-

mum of the transient response. With a small amount of practice these values of T can be determined with one to three iterations per value. The values of T_p determined from Equation 16 can then be inserted in Equation 17 to determine which maximum gives the peak overshoot, and the amount of that overshoot. From the foregoing we see that the calculations for time and magnitude of peak overshoot are inseparable. Each must be known before the other can be determined correctly. This is not a severe handicap, however, as the peak overshoot occurs on the first maximum for practically all combinations of n_1 and n_2 which include values of n_2 less than two.

Loci of constant values of magnitude of peak overshoot (M_p) are shown on the z-plane in Figure V-3. Another set of loci of constant values of peak overshoot with a compressed ordinate scale to permit extension of the plot to include larger values of n_2 is shown in Figure V-4. Although Figure V-4 does not represent a complex plane due to the difference in scales used on the axes, it may be considered as one for the purpose of reading values of peak overshoot associated with values of n_1 and n_2.

Loci of constant values of artificial time of peak overshoot (T_p) are shown on the z-plane in Figure V-5. An interesting point in connection with T_p is that as n_2 increases the peak overshoot occurs on the second or subsequent maximum of the transient response curve, and the

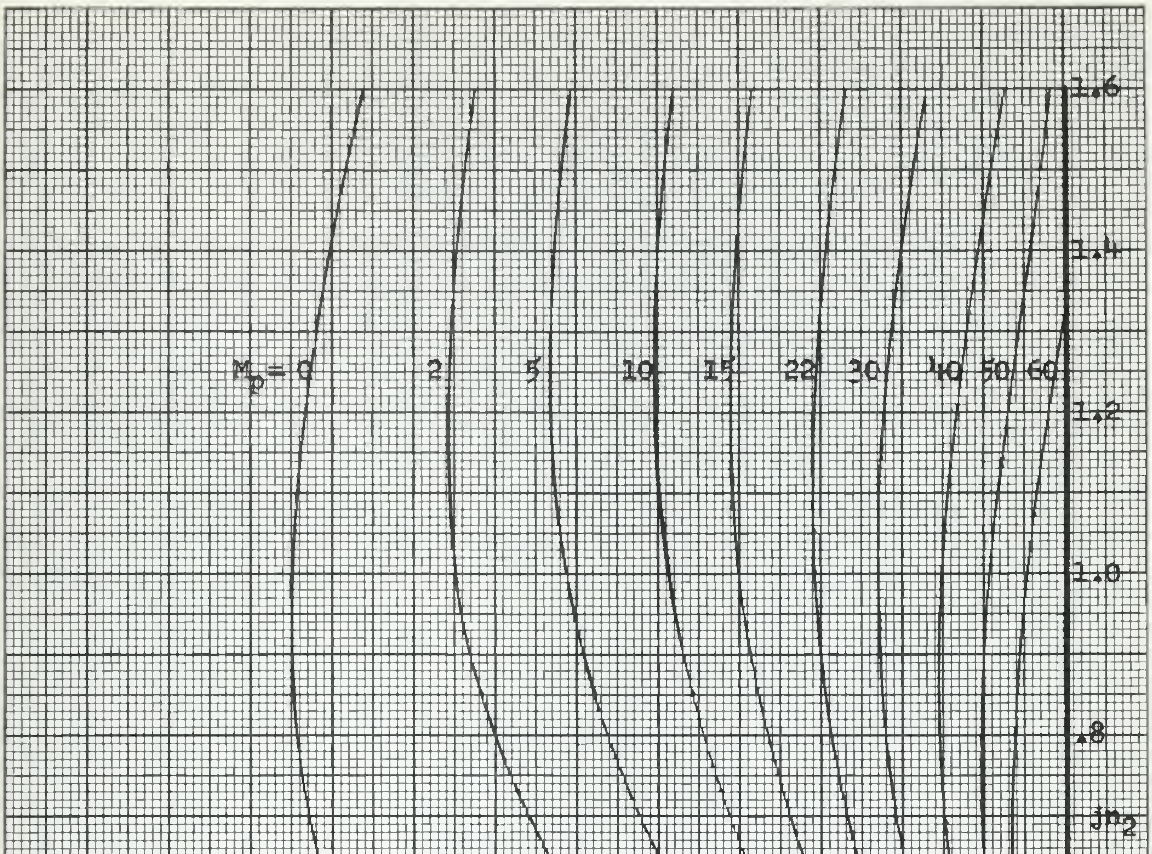

FIGURE V-3. PEAK OVERSHOOT (PERCENT) PURE THIRD ORDER SYSTEM

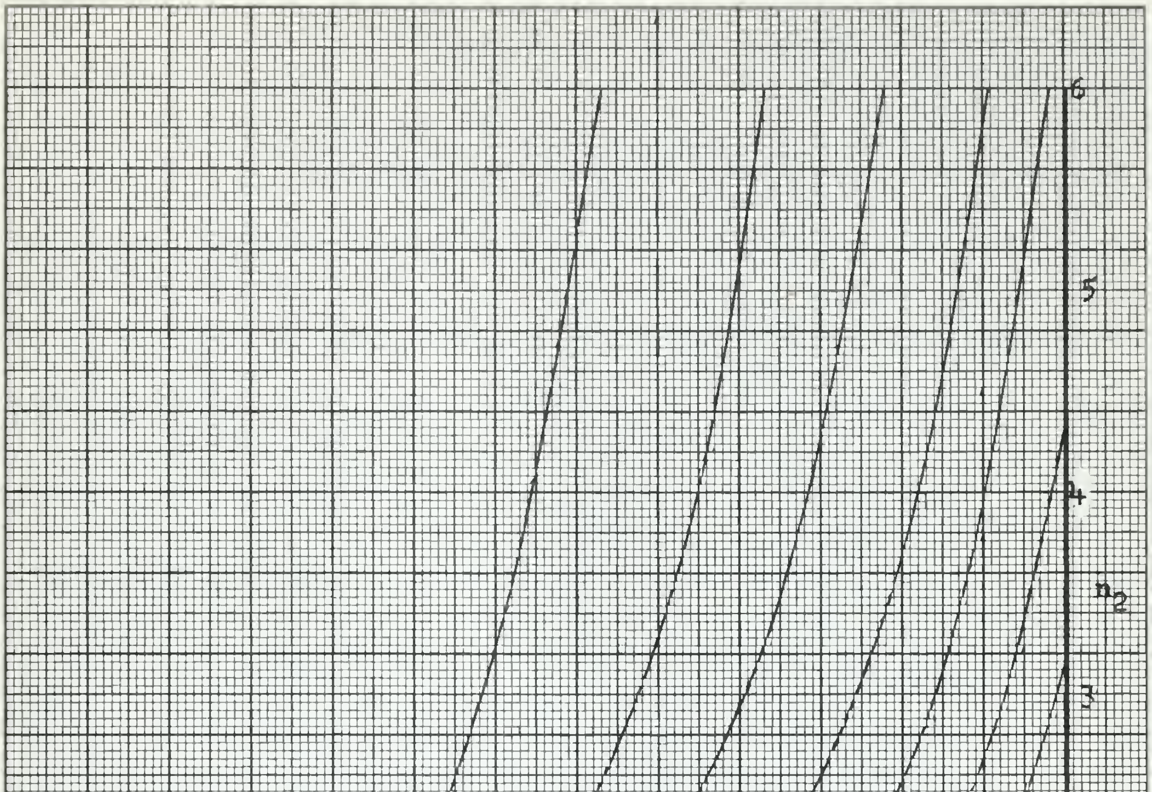

FIGURE V-4
MAGNITUDE OF PEAK OVERSHOOT (PERCENT)
PURE THIRD ORDER SYSTEM

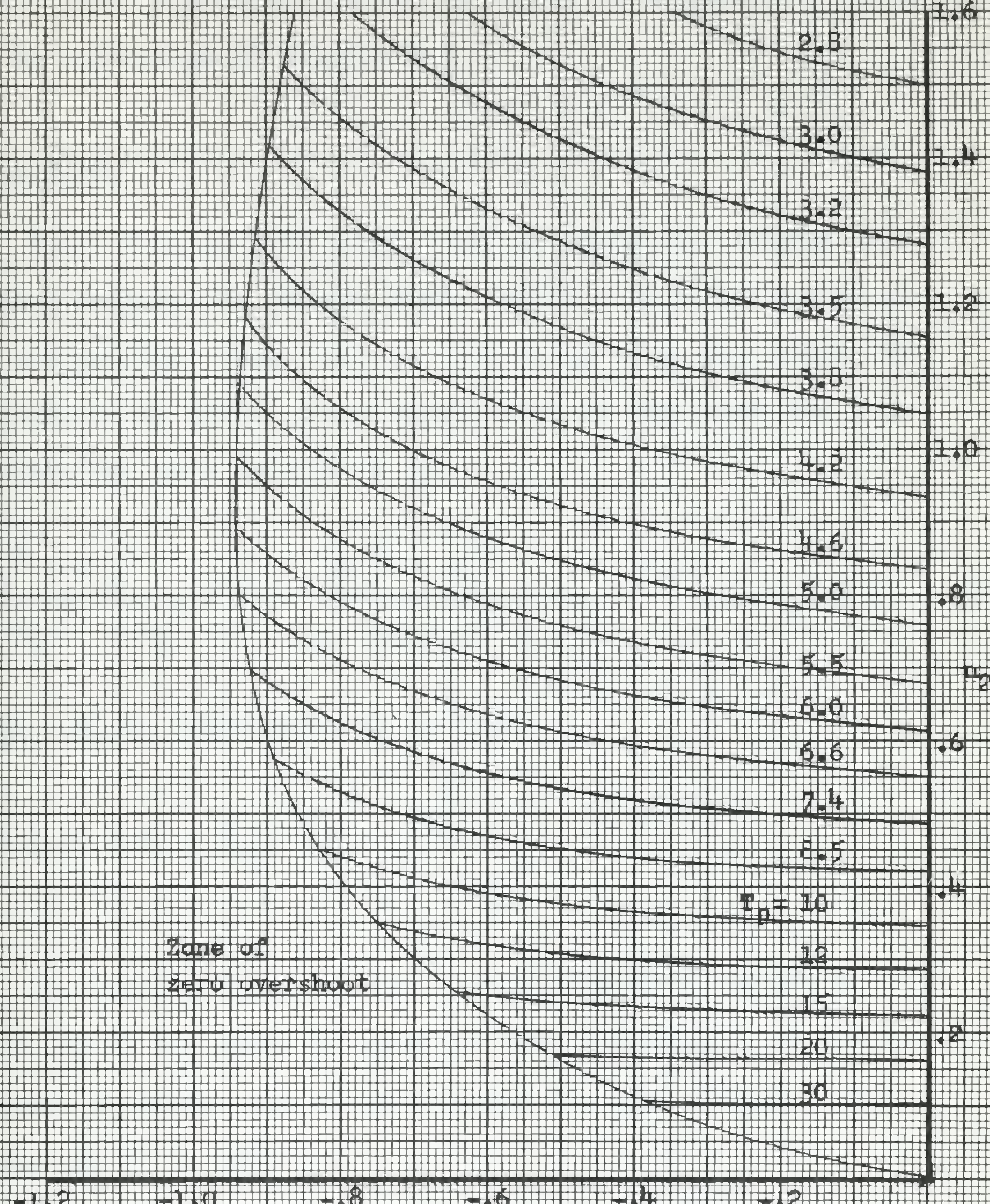

FIGURE V-5. ARTIFICIAL TIME OF PEAK
OVERSHOOT - PURE THIRD ORDER SYSTEM

time of peak overshoot is no longer unique. The zone in which this situation exists is shown in Figure V-6 which is an extension of Figure V-5.

Rise Time

Rise time is the first time at which the output of a system equals its steady state value. This is equivalent to saying that rise time is the first time for which the overshoot is zero. From Equation 5 we see that the condition for zero overshoot is

$$\frac{n_1^2 + n_2^2}{(1-n_1)^2 + n_2^2} e^{-T} = \frac{\sqrt{n_1^2 + n_2^2}}{n_2 \sqrt{(1-n_1)^2 + n_2^2}} e^{-n_1 T} \sin(n_2 T - \theta - \emptyset) \qquad (18)$$

This can be rearranged as follows

$$\sin(n_2 T - \theta - \emptyset) = n_2 \sqrt{\frac{n_1^2 + n_2^2}{(1-n_1)^2 + n_2^2}}\, e^{-(1-n_1)T} \qquad (19)$$

Obviously, this too is a transcendental equation, but just as in the case of time of peak overshoot, it can be solved by trial and error methods for any selected combination of n_1 and n_2.

Loci of constant values of artificial rise time (T_r) are shown on the z-plane in Figure V-7.

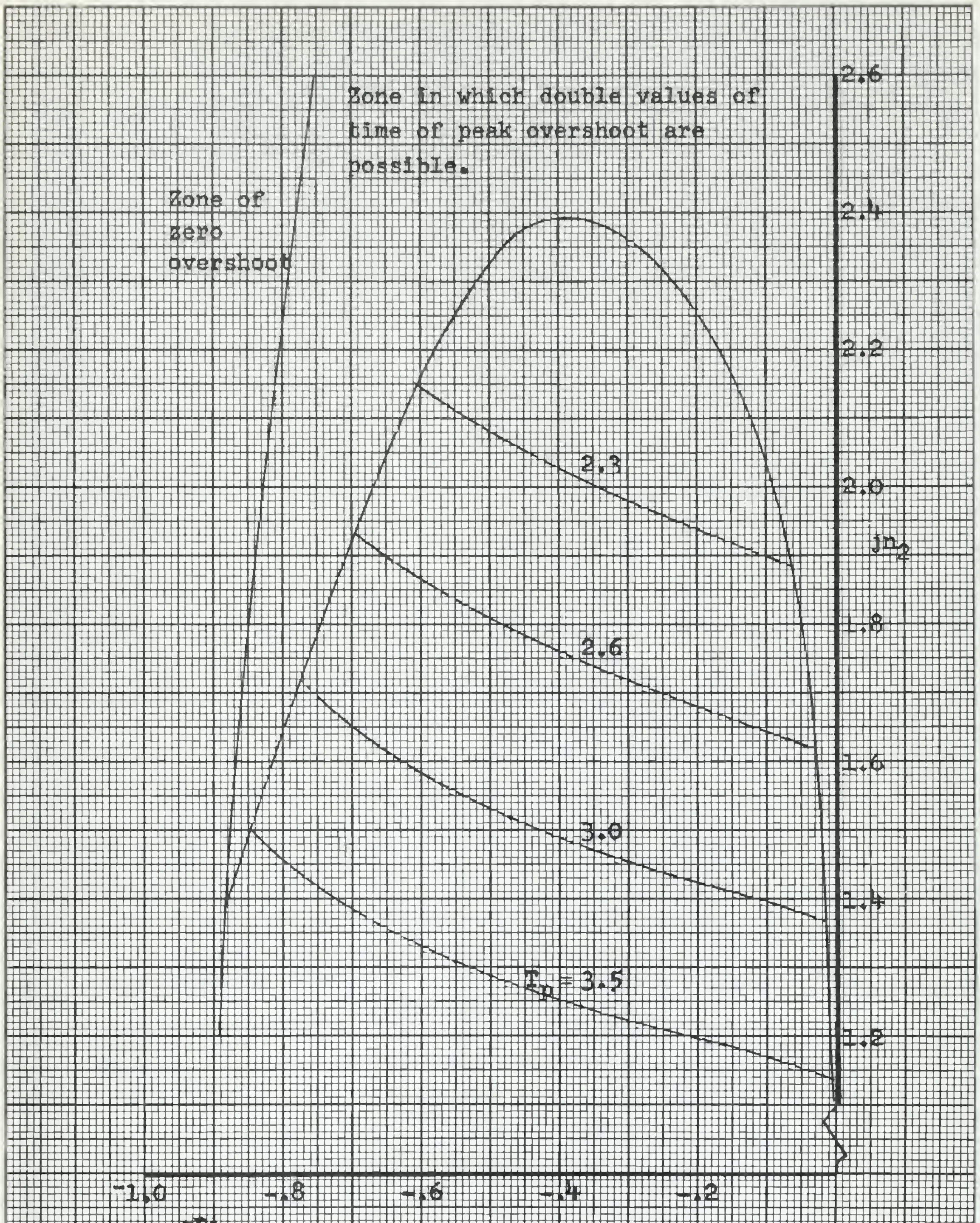

FIGURE V-6. ZONE OF DOUBLE VALUES OF TIME OF PEAK OVERSHOOT - PURE THIRD ORDER SYSTEM

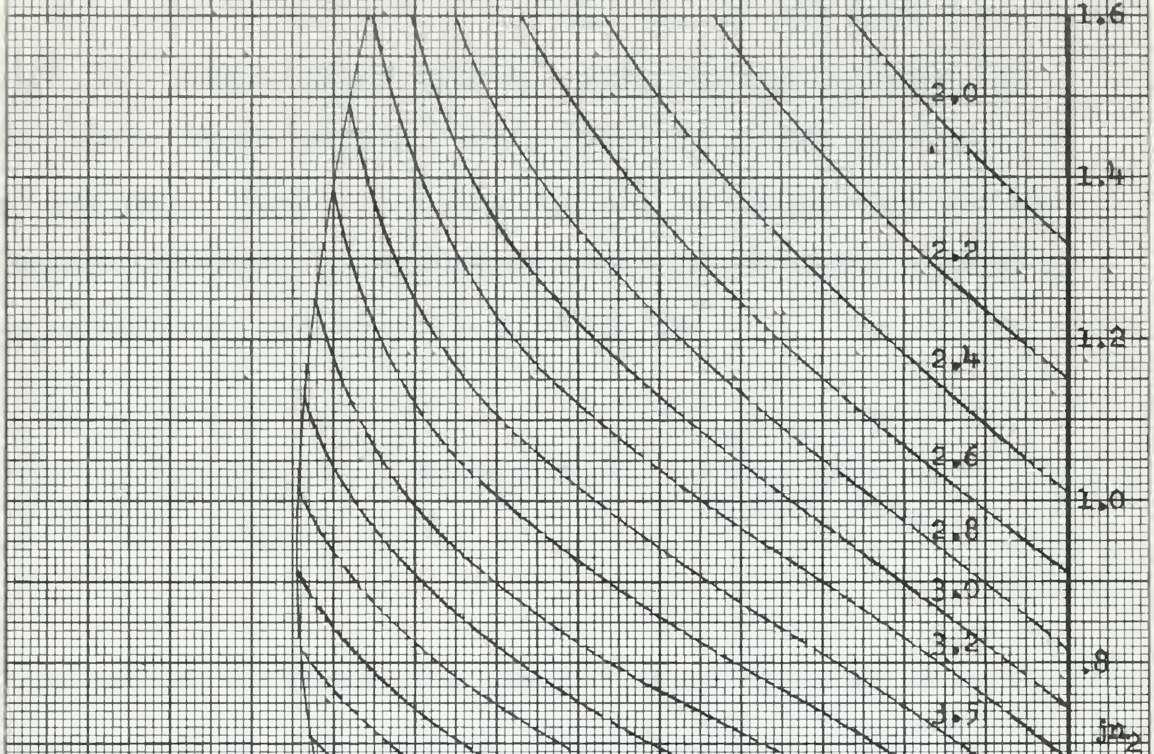

FIGURE V-7. ARTIFICIAL RISE TIME
PURE THIRD ORDER SYSTEM

Settling Time

As normally defined, settling time is the last time at which the output error exceeds some fixed percentage of the steady state output. This error is often specified as two percent. Using this definition, settling time is extremely difficult to determine for a third order system as the condition that must exist at T_s is

$$\frac{\sqrt{n_1^2 + n_2^2}}{n_2\sqrt{(1-n_1)^2 + n_2^2}}\, e^{-n_1T} \sin(n_2T-\theta-\emptyset) - \frac{n_1^2 + n_2^2}{(1-n_1)^2 + n_2^2}\, e^{-T} = \pm .02 \tag{20}$$

To determine settling time exactly would require several trial and error solutions of Equation 20 or a plot of Equation 5 for each combination of n_1 and n_2. The labor involved in either method is prohibitive.

The settling time for the third order system as defined here is the time at which the sum of (1 - exponential term) plus the height of the envelope of the sinusoidal term is equal to two percent of the steady state output. The equation to be satisfied for this condition is

$$\frac{\sqrt{n_1^2 + n_2^2}}{n_2\sqrt{(1-n_1)^2 + n_2^2}}\, e^{-n_1T_s} + \frac{n_1^2 + n_2^2}{(1-n_1)^2 + n_2^2}\, e^{-T_s} = .02 \tag{21}$$

Loci of constant values of artificial settling time (T_s) are shown on the z-plane in Figure V-8. Settling time

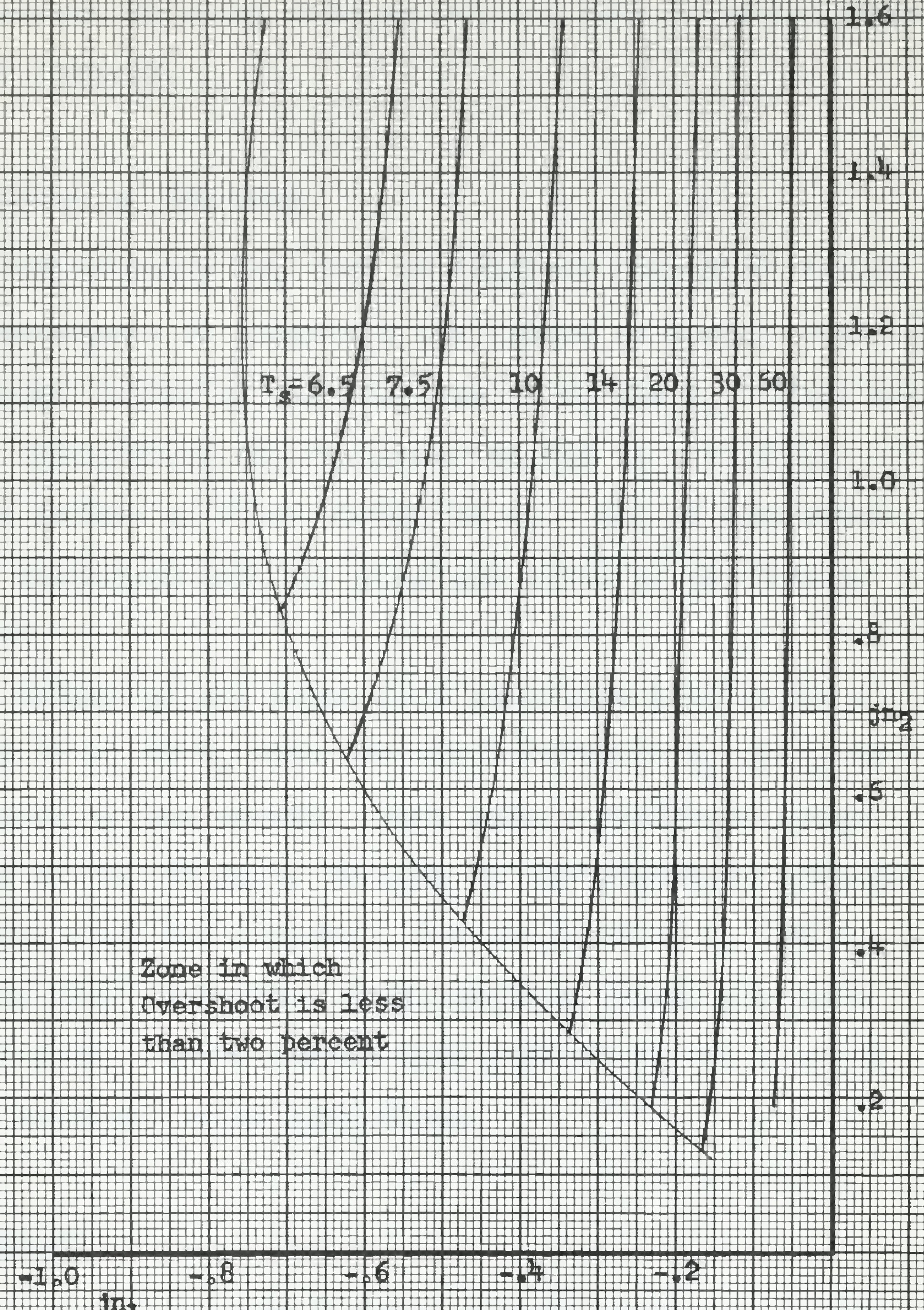

FIGURE V-8. ARTIFICIAL SETTLING TIME
PURE THIRD ORDER SYSTEM

for systems with less than two percent overshoot is not shown because settling time as defined here is meaningless for such systems.

Construction of the Loci

In constructing the loci of constant parameter values it was necessary to calculate the values of each parameter for a large number of combinations of n_1 and n_2. This was done by calculating the parameter values as a function of n_2 for a number of fixed values of n_1. All of the calculations were performed by trial and error as previously indicated.

The values of the parameters were then plotted as a function of n_1 or n_2 with the other variable n_1 or n_2 as a parameter of the curves. This gave a series of cross-sections of the surfaces defined by the transient response parameters over the z-plane. The coordinates n_1 and n_2 were then read from these cross-section curves for each locus desired. These points were plotted on the z-plane and the constant value loci drawn.

Observations

The loci of constant values of peak overshoot become straight lines as they approach the origin of the complex plane. This is consistent with the usual second order

approximation for the case in which the value of the real pole is several times the value of the undamped natural frequency of the system. In Figure V-9 the validity of this approximation is demonstrated by a comparison of the peak overshoot of the pure second order system with that of the pure third order system for small values of n_1 and n_2.

Another area of interest on the plot of overshoot loci is the area of zero overshoot. The zero overshoot locus represents the boundary where the magnitude of peak overshoot becomes too small to calculate. The boundary as shown represents a value of the order of e^{-10}. For values of n_1 greater than about 0.94 the peak overshoot is too small to calculate, regardless of the value of n_2. In this case the response is essentially that of a first order system with a decaying ripple superimposed on the rising exponential curve.

As noted previously, there exists a region in which the time of peak overshoot is not necessarily unique. This area is shown in Figure V-6. The zone is bounded by a line through the values of n_1 and n_2 for which the height of the first and second maximums of the transient response curve are identical. Obviously this includes the entire imaginary axis.

For small values of n_2 the time of peak overshoot is relatively independent of the value of n_1. In fact, the

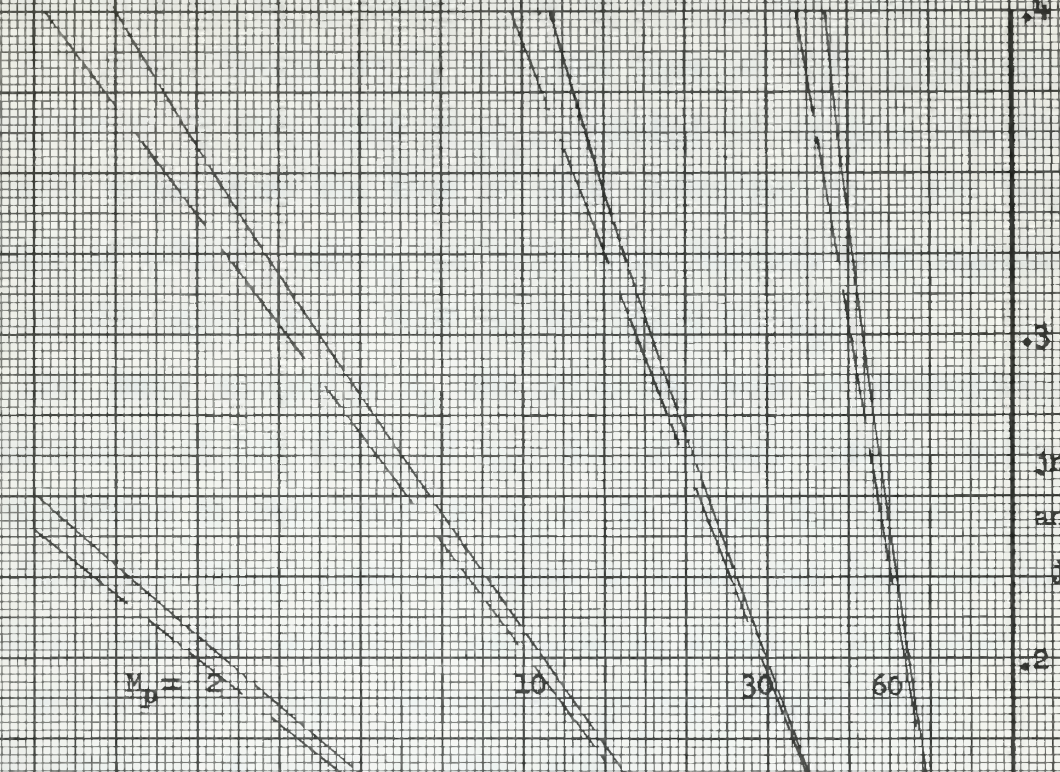

FIGURE V-9. MAGNITUDE OF PEAK OVERSHOOT (PERCENT)
PURE SECOND ORDER AND PURE THIRD ORDER SYSTEMS

only area where time of peak overshoot becomes sensitive to the value of n_1 is where both n_1 and n_2 are large.

For small values of n_2 the rise time is also relatively independent of the value of n_1, although it is more dependent on this value than the time of peak overshoot is. For large values of both n_1 and n_2 the rise time becomes more dependent on the value of n_1 than on the value of n_2.

For small values of n_1 the settling time is essentially independent of the value of n_2, and settling time is never as dependent on the value of n_2 as it is on the value of n_1.

To avoid heavy concentrations of lines the loci are not plotted near the origin of the complex plane. This does not detract from their usefulness as the second order approximation is valid in this area.

<u>Use of the Curves</u>

The curves for the pure third order system are used in exactly the same manner as the curves of Chapters III and IV. Since their use has already been explained there, the explanation will not be repeated here. Instead an actual example is presented.

Assume that a system has the following closed loop performance function

$$PF(s) = \frac{2500}{(s+5)(s+2+j4)(s+2-j4)}$$

Dividing the real and imaginary parts of the complex poles by the value of the real pole gives:

$$n_1 = \frac{2}{5} = .4$$

$$n_2 = \frac{4}{5} = .8$$

Entering Figures V-3, V-5, V-7, and V-8 with these values of n_1 and n_2 we can read the following parameter values:

$$M_p = 13\%$$

$$T_p = 5.1$$

$$T_r = 3.8$$

$$T_s = 10.2$$

Dividing the three artificial times by the value of the real pole gives the real times:

$$t_p = \frac{5.1}{5} = 1.02 \text{ sec.}$$

$$t_r = \frac{3.8}{5} = .76 \text{ sec.}$$

$$t_s = \frac{10.2}{5} = 2.04 \text{ sec.}$$

From the performance function we see that the transient oscillating frequency is four radians per second, or the period is $\frac{\pi}{2}$ seconds. The three values of time, the period of oscillations, magnitude of peak overshoot and the fact that the slope is zero at zero time and at time of peak overshoot constitute sufficient information from which to

sketch the transient response. The transient response is sketched in Figure V-10.

Now if we apply a unit step input to the system and obtain the transient response by performing the inverse transformation, we get

$$C(t) = 1 - 8e^{-5t} + 1.116e^{-2t}\sin(229t - 169.7)^{\circ}$$

This output has been calculated and is plotted as a function of time in Figure V-11. From an expanded scale plot of this output, transient response parameters have been determined. These are compared with the predicted values below.

Parameter	Predicted Value	Value from Transient Response
t_p	1.02 sec.	1.025 sec.
t_r	.76 sec.	.76 sec.
t_s	2.04 sec.	2.00 sec.
M_p	13 %	12.7 %

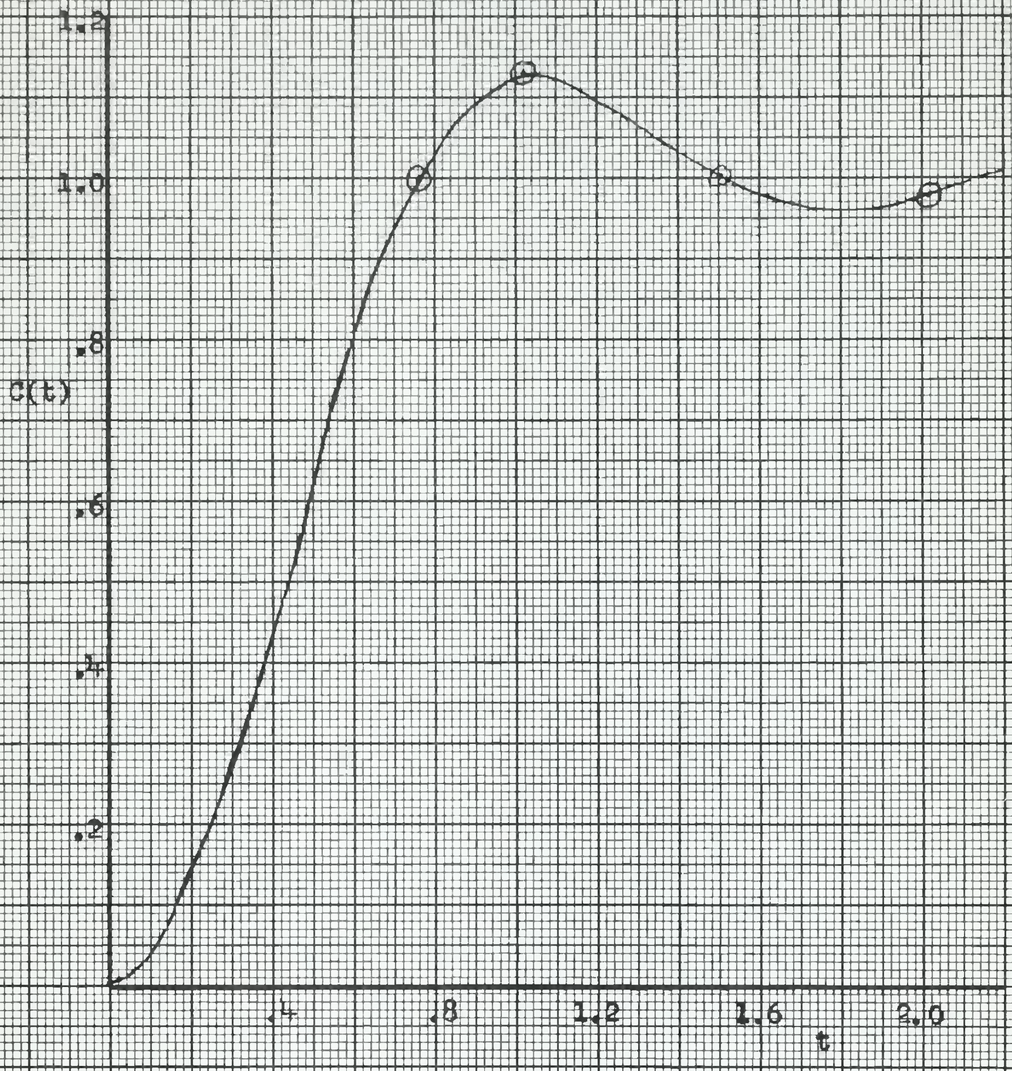

FIGURE V-10
APPROXIMATE TRANSIENT RESPONSE
PURE THIRD ORDER EXAMPLE

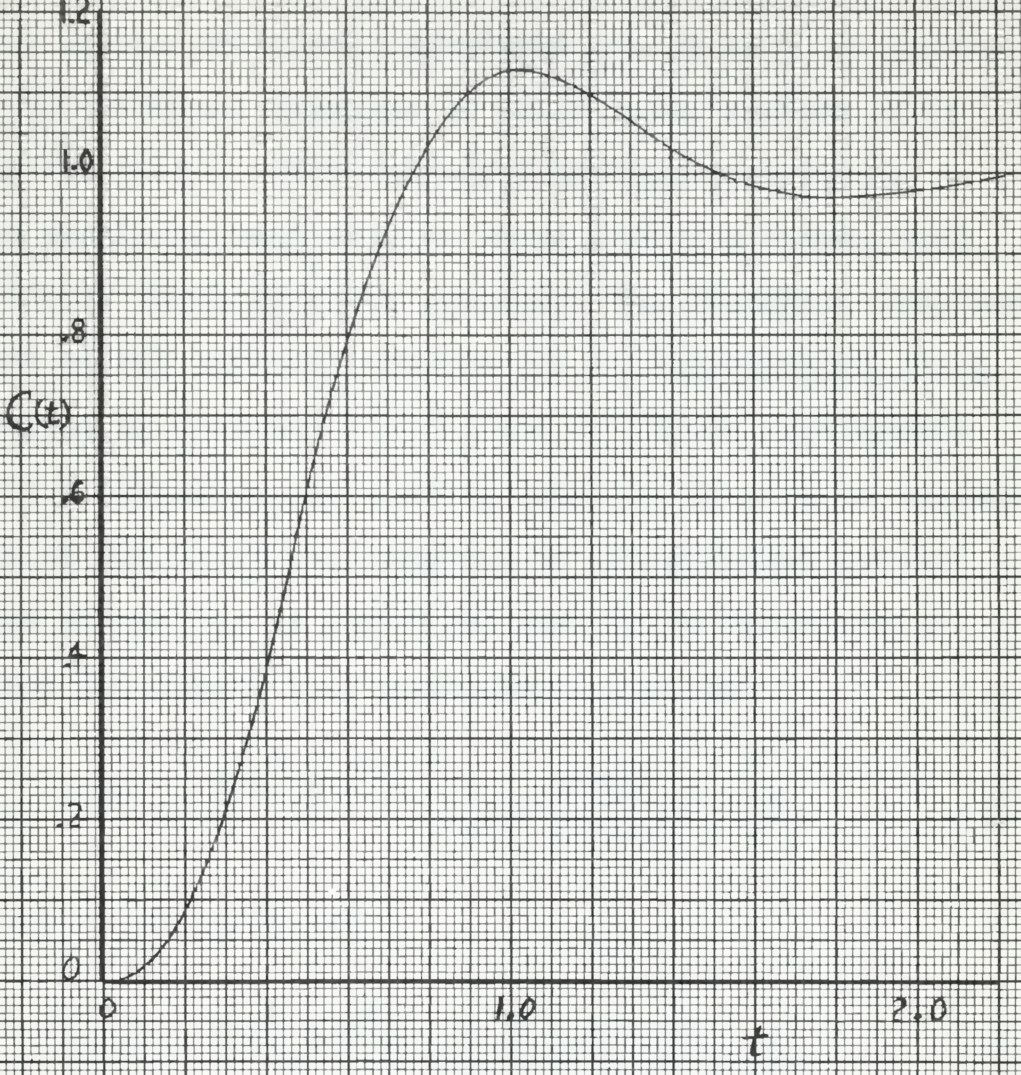

FIGURE V-11

ACTUAL TRANSIENT RESPONSE
PURE THIRD ORDER EXAMPLE

Chapter VI

EXTENSION OF THE TECHNIQUE

The curves previously developed can be used to approximate the transient response of third order systems with one or two zeros. This extension is made possible by a simple algebraic manipulation which breaks a system down into two or three parallel systems. The transient response of each of the parallel systems can be approximated using the curves of Chapters III, IV and V. These approximate transient responses can then be combined to give an approximation of the transient response of the original system.

The scheme can be explained best by an illustration. Consider the closed loop performance function

$$PF(s) = \frac{c(a^2 + b^2)}{d} \cdot \frac{(s+d)}{(s+c)\left[(s+a)^2 + b^2\right]} \tag{1}$$

This can be rewritten as

$$PF(s) = \frac{c(a^2 + b^2)}{d} \cdot \frac{(s + c + f)}{(s+c)\left[(s+a)^2 + b^2\right]} \tag{2}$$

where f is defined by

$$f = (d - c) \tag{3}$$

We can then write the performance function as

$$PF(s) = \frac{c}{d}\left[\frac{a^2 + b^2}{(s+a)^2 + b^2}\right] + \frac{f}{d}\left[\frac{(a^2 + b^2)c}{(s+c)\left[(s+a)^2 + b^2\right]}\right] \quad (4)$$

Applying a unit step input and taking the inverse transform gives the transient response as

$$C(t) = \frac{c}{d}\left[\begin{matrix}\text{Pure second}\\ \text{order response}\end{matrix}\right] + \frac{f}{d}\left[\begin{matrix}\text{Pure third}\\ \text{order response}\end{matrix}\right]$$

To sketch this transient response the pure second order response and the pure third order response must both be sketched first by the methods previously described. Both must be sketched as functions of real time. At any instant the total response is $\frac{c}{d}$ multiplied by the value of the pure second order response at that time plus $\frac{f}{d}$ multiplied by the value of the pure third order response at the same time. The two responses can be combined most rapidly by graphical methods.

As an example, consider the performance function

$$PF(s) = \frac{26(s+2)}{(s+4)\left[(s+2)^2 + 3^2\right]} \quad (5)$$

In this case $f = (d - c) = (2 - 4) = -2$. Therefore

$$PF(s) = \frac{4}{2}\left[\frac{13}{(s+2)^2 + 3^2}\right] - \frac{2}{2}\left[\frac{52}{(s+4)\left[(s+2)^2 + 3^2\right]}\right] \quad (6)$$

$$C(t) = 2\left[\begin{matrix}\text{Pure second}\\ \text{order response}\end{matrix}\right] -1\left[\begin{matrix}\text{Pure third}\\ \text{order response}\end{matrix}\right]$$

For the pure second order response, $a = 2$ and $b = 3$. From the curves in Chapter III values of parameters are:

$$T_p = 2.1$$

$$T_r = 1.47$$

$$T_s = 4.1$$

$$M_p = 12.8\%$$

Dividing the artificial times by the value of a yields:

$$t_p = 1.05 \text{ sec.}$$

$$t_r = .735 \text{ sec.}$$

$$t_s = 2.05 \text{ sec.}$$

For the third order response:

$$n_1 = \frac{2}{4} = .5$$

$$n_2 = \frac{3}{4} = .75$$

From the curves in Chapter V values of parameters are:

$$T_p = 5.6$$

$$T_r = 4.3$$

$$T_s = 8.5$$

$$M_p = 6.5\%$$

Dividing these artificial times by the value of c yields:

$$t_p = 1.4 \quad \text{sec.}$$

$$t_r = 1.075 \quad \text{sec.}$$

$$t_s = 2.125 \quad \text{sec.}$$

Note that in both cases the transient oscillating frequency is $\frac{3}{2\pi}$ cycles per second.

The process of combining the two responses is accomplished graphically. This is illustrated in Figure VI-1 where the two transient response curves and their sum have been sketched.

The same method can be used to approximate the response of third order systems with two zeros. This merely involves more labor and of course the accuracy declines with the addition of zeros.

Consider the following closed loop performance function

$$PF(s) = \frac{c(a^2 + b^2)}{d_1 d_2} \cdot \frac{(s+d_1)(s+d_2)}{(s+c)\left[(s+a)^2 + b^2\right]} \tag{7}$$

This can be rewritten as

$$PF(s) = \frac{c(a^2 + b^2)}{d_1 d_2} \cdot \frac{(s+c+f_1)(s+c+f_2)}{(s+c)\left[(s+a)^2 + b^2\right]} \tag{8}$$

where

$$f_1 = d_1 - c$$

$$f_2 = d_2 - c$$

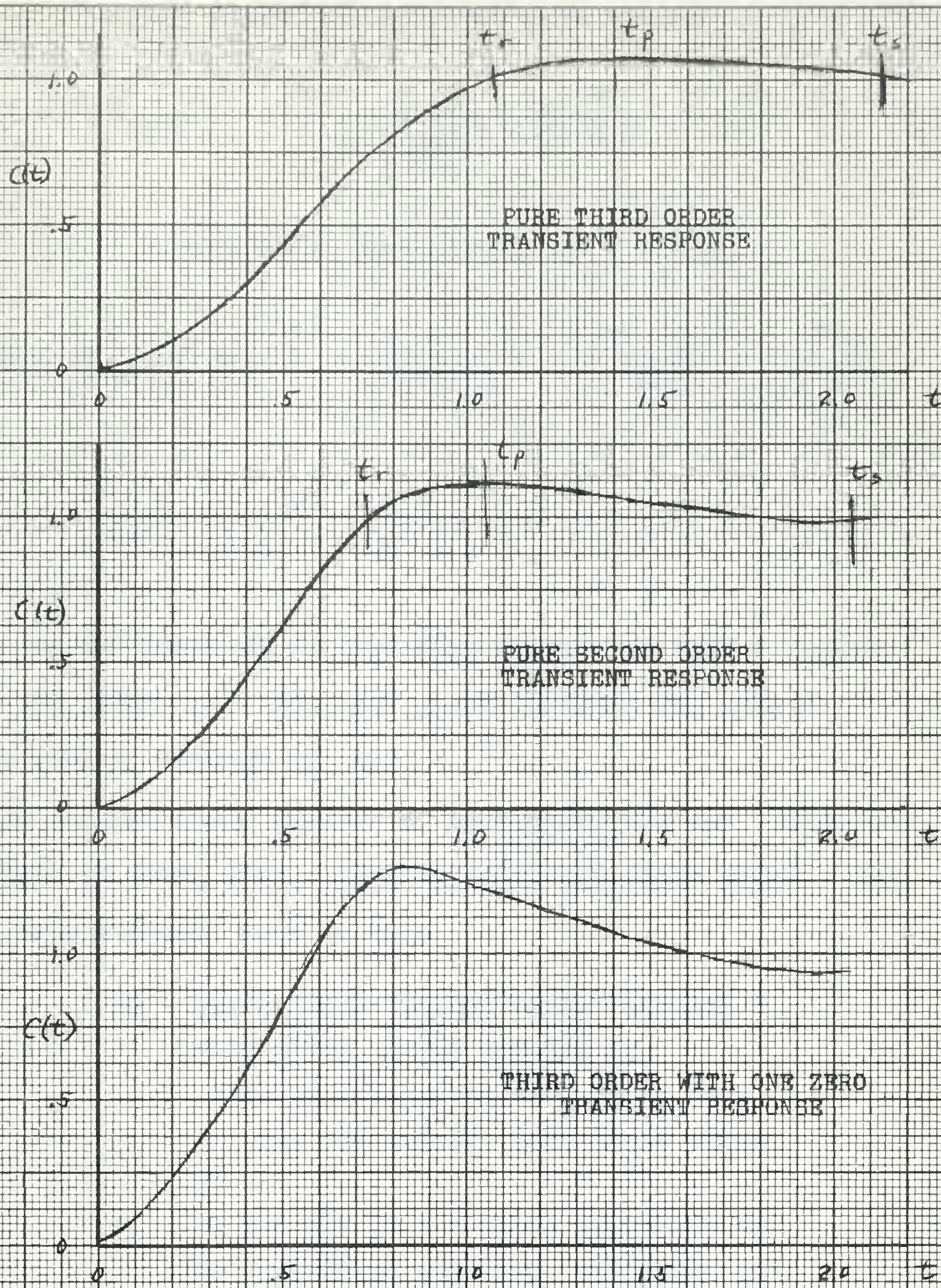

FIGURE VI-1

TRANSIENT RESPONSE APPROXIMATION
THIRD ORDER SYSTEM WITH ONE ZERO

Multiplying the terms in the numerator yields

$$PF(s) = \frac{c(a^2 + b^2)}{d_1 d_2} \cdot \frac{(s+c)^2 + (f_1 + f_2)(s+c) + f_1 f_2}{(s+c)\left[(s+a)^2 + b^2\right]} \tag{9}$$

This can be broken into three terms as follows:

$$PF(s) = \frac{c^2}{d_1 d_2} \cdot \frac{a^2 + b^2}{c} \cdot \frac{s+c}{(s+a)^2 + b^2} + \frac{(f_1 + f_2)c}{d_1 d_2} \cdot \frac{a^2 + b^2}{(s+a)^2 + b^2}$$
$$+ \frac{f_1 f_2}{d_1 d_2} \cdot \frac{c(a^2 + b^2)}{(s+c)\left[(s+a)^2 + b^2\right]} \tag{10}$$

This can be sketched in the same manner as before except that the transient response curve will be the sum of three curves instead of two.

No satisfactory method of approximating the transient response of more complicated systems has been found. The only alternatives seem to be:

a. Approximate higher order systems with lower order systems by ignoring the poles and zeros that are far out in the left hand plane, or

b. Expand the performance function by the method of partial fractions into fragments of manageable size, utilizing the techniques developed in this thesis whenever possible.

Chapter VII

CONCLUSIONS AND RECOMMENDATIONS

The servomechanisms investigated in this thesis have been divided into classes according to the number and type of poles and zeros of their closed loop performance functions. This was done to facilitate the investigation of the transient responses of the different classes of systems. Systems investigated in this thesis are those whose closed loop performaance functions contain the following pole-zero configurations:

a. Two complex poles only. (Pure second order systems)

b. Two complex poles and one zero. (Second order system with one zero)

c. Two complex poles and one real pole. (Pure third order system)

It has been shown that for each of these classes of systems the transient response of systems with the same relative pole-zero configuration can be nondimensionalized. The transient response can be expressed as a function of an artificial time and two complex coordinates. For the

pure second order system the artificial time is the magnitude of the real part of the complex poles multiplied by real time and the complex coordinates are the real and imaginary parts of the complex poles. For systems with three singularities the artificial time is the magnitude of the real pole or zero multiplied by real time and the complex coordinates are the real and imaginary parts of the complex poles divided by the magnitude of the real pole or zero.

With the transient response expressed in this manner, magnitude and time parameters of the response can be expressed in terms of percentage of steady state output and artificial time respectively. Loci of constant values of four transient response parameters for each of the three classes of systems listed above have been mapped onto a complex plane. These parameters are magnitude and artificial time of peak overshoot, artificial rise time and artificial settling time. For the pure second order system the loci are mapped directly onto the s-plane, while for the systems with three singularities the loci are mapped onto an auxiliary complex plane.

The approximate value of these four parameters for practically any specific system of the three classes discussed can be determined from these families of loci of constant parameter values. A rather accurate transient

response curve can then be sketched using these four parameters, transient oscillating frequency, initial slope and slope at time of peak overshoot. This has been demonstrated by example.

Although the discussion in this thesis has been limited to systems which exhibit unity feedback characteristics, this is not a necessary restriction. The closed loop performance function of a non-unity feedback system can be written as a constant multiplied by the performance function of a unity feedback system. Regardless of the value of this constant, it will appear in the transient response as a multiplier of all terms. Therefore, the value of the constant will not effect the value of any of the transient response parameters determined in this thesis. The transient response of such a system will merely have a steady state value different from unity.

Although loci of only four transient response parameters have been mapped onto complex planes in this thesis, loci of other parameters could be mapped in a similar manner. Other such parameters which would appear to be useful in improving the resulting transient response sketches include:

a. Delay Time (Time at which the output first equals one-half the steady state value).

b. Time and Magnitude of First Undershoot.

Inclusion of additional singularities in the closed loop performance funtion of a system presents additional problems. Although the technique developed in this thesis can be applied to such systems, one set of loci is not sufficient for all systems of a class. Consider a system with two complex and two real singularities. Application of the mapping technique would require fixing one of the real variables as a function of the other real variable and construction of a set of loci for this fixed relative configuration of the two variables. Additional families of loci would be required for each fixed configuration of these two real variables for which knowledge of the parameters was desired.

One method of avoiding the additional mapping required for third order systems with one or two zeros has been presented. This involves separation of such systems into parallel components whose individual responses can be approximated using the parameter loci presented in this thesis. Once the component responses have been approximated they may be combined to give a resultant approximation of the system transient response. Using this scheme the accuracy of the resultant approximation is considerably less than the accuracy of the component

approximations. However, careful construction of the component approximations should permit the construction of a useful resultant approximation, at least for the third order system with one zero.

The technique developed in this thesis is not restricted to the mapping of transient response parameters. Frequency response parameters could also be mapped onto a complex plane by a similar technique.

The authors acknowledge that they have not exhausted the possibilities of the technique presented in this thesis. In fact, they believe that there are a vast number of possibilities yet to be investigated. However, due to the labor involved in the mapping technique, further investigations should be conducted with the aid of a digital computer. Specific areas recommended for investigation using a computer are:

a. The three classes shown in this thesis to extend the loci plots to include larger values of the complex coordinates.

b. The third order system with one zero to include construction of sets of parameter loci for each of a number of fixed relative configurations of the real pole and zero.

c. Construction of loci of additional parameters such as delay time and time and magnitude of the first undershoot for the systems listed in a and b.

d. Construction of loci of constant values of frequency response parameters.

e. The possibility of mapping statistical quantities such as integral error squared for various inputs.

BIBLIOGRAPHY

1. Chu, Yaohan: Synthesis of Feedback Control Systems by Phase Angle Loci. Transactions A.I.E.E., part II, Applications and Industry, Vol. 71, 1952.

2. Brown, G.S. and Campbell, D.P.: Principles of Servomechanisms, Chapter 11, New York: John Wiley and Sons, Inc., 1948.

3. Thaler, George J. and Brown, Robert G.: Analysis and Design of Feedback Control Systems, New York: McGraw Hill Book Co., Inc., 1960.

4. Truxal, John G.: Automatic Feedback Control System Synthesis, New York: McGraw Hill Book Co., Inc., 1955.

5. Liu, Y.J.: Servomechanisms: Copyright 1941 by Massachusetts Institute of Technology, Cambridge, Mass.

6. Mitrovic, Dusan: Graphical Analysis and Synthesis of Feedback Control Systems: Part I, Theory and Analysis; Part II, Synthesis: Transactions A.I.E.E., pt. II, January 1959.

7. Gardner, Murray F. and Barnes, John L.: Transients in Linear Systems, Volume I, New York: John Wiley and Sons, Inc., 1942.

8. Wylie, C.R., Jr.: Advanced Engineering Mathematics, New York: McGraw Hill Book Co., Inc., 1951.